Cross-bispectra constraints on modified gravity theories from the Nancy Grace Roman Space Telescope and the Rubin Observatory Legacy Survey of Space and Time

Chen Heinrich* and Olivier Doré

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA and California Institute of Technology, Pasadena, California 91109, USA

(Received 19 June 2020; accepted 10 November 2020; published 31 December 2020)

One major goal of upcoming large-scale-structure surveys is to constrain dark energy and modified gravity theories. In particular, galaxy clustering and gravitational lensing convergence are probes sensitive to modifications of general relativity. While the standard analysis for these surveys typically includes power spectra or 2-point correlation functions, it is known that the bispectrum contains additional information that could offer improved constraints on parameters when combined with the power spectra. However, the use of bispectra has been limited so far to one single probe, e.g., the lensing convergence bispectrum or the galaxy bispectrum. In this paper, we extend the formalism to explore the power of cross-bispectra between different probes, and exploit their ability to break parameter degeneracies and improve constraints. We study this on a test case of lensing convergence and galaxy density auto- and cross-bispectra, for a particular subclass of Horndeski theories parametrized by the running of the Planck mass c_M and the braiding parameter c_B. Using the 2000 deg^2 notional survey of the Nancy Grace Roman Space Telescope with overlapping photometry from the Rubin Observatory Legacy Survey of Space and Time, we find that a joint power spectra and bispectra analysis with three redshift bins at $l_{\max} = 1000$ yields $\sigma_{c_M} = 1.0$ and $\sigma_{c_B} = 0.3$, both a factor of $\sim$1.2 better than the power spectra results; this would be further improved to $\sigma_{c_M} = 0.7$ and $\sigma_{c_B} = 0.2$ if $l_{\max} = 3000$ is taken. Furthermore, we find that using all possible cross-bispectra between the two probes in different tomographic bins improves upon auto-bispectra results by a factor of 1.3 in σ_{c_M}, 1.1 in σ_{c_B}, and 1.3 in σ_{Ω_m}. We expect that similar benefits of using cross-bispectra between probes could apply to other science cases and surveys.

DOI: 10.1103/PhysRevD.102.123549

I. INTRODUCTION

Recent measurements have shown that the Universe is consistent with a ΛCDM model, exhibiting an epoch of accelerated expansion of the Universe. The observation of the accelerated epoch poses some theoretical challenges. Three possibilities are typically considered: (1) a cosmological constant; (2) a scalar field giving rise to a possibly evolving energy density, dubbed dark energy; and (3) the theory that general relativity (GR) does not hold on cosmological scales, requiring modified gravity theories (MG). While GR is a well-tested theory on some scales such as the solar system scale, whether it can be successfully extrapolated to all cosmological scales over many orders of magnitude is still an assumption that remains to be tested. Cosmology holds great promise for testing alternative theories of gravity as they would leave distinguishable signatures on probes such as the clustering of galaxies, gravitational lensing, and redshift space distortions.

Upcoming Stage-IV large-scale-structure surveys such as EUCLID[1] [1], Nancy Grace Roman Space Telescope[2] [2], Rubin Observatory Legacy Survey of Space and Time (LSST)[3] [3], and DESI[4] [4] are optimally designed to maximize constraints on dark energy and modified gravity. By detecting hundreds of millions of galaxies in large areas of the sky in imaging and spectroscopy, these surveys allow us to construct powerful statistical probes that can distinguish between different theories. The Roman Space Telescope is a particularly promising survey that would notionally map a 2000 deg^2 area of the sky in depth in both spectroscopy and imaging, enabling multiprobe analysis as well as exquisite systematics control [2,5–7].

To take advantage of these superb capabilities of future surveys, it is no longer sufficient to restrict ourselves to the

*chenhe@caltech.edu

[1]www.euclid-ec.org.
[2]Formerly the Wide Field Infrared Survey Telescope (WFIRST): https://roman.gsfc.nasa.gov/.
[3]https://www.lsst.org/.
[4]http://desi.lbl.gov/.

typical analysis using power spectra or 2-pt correlation functions. Additional information is known to exist in higher-order statistics such as the bispectrum. Combining bispectrum with power spectrum observations typically leads to improved constraints on parameters. In this paper, we explore how adding bispectra, and in particular cross-bispectra measurements between two probes, provide improved constraints on modified gravity theories for the overlapping 2000 deg^2 between the Roman Space Telescope and the LSST survey.

Just as using cross-power spectra can break parameter degeneracies leading to improved constraints, we expect a similar effect by cross-correlating probes in the bispectra. Typically the lensing convergence bispectrum and the galaxy bispectrum have been studied individually (e.g., [8–11]). Here, we combine for the first time the galaxy and lensing convergence auto-bispectra as well as their cross-bispectra between these probes in different tomographic bins, to exploit their potential to improve on parameter constraints. We demonstrate such improvement on a particular subclass of Horndeski models for the Roman Space Telescope, but expect similar benefits to extend potentially to other science cases or other surveys.

The paper is structured as follows. In Sec. II, we present the background on Horndeski theories and define the subclass of MG models that we study. In Secs. III and IV, respectively, we describe the modeling of the power spectrum and bispectrum observables, as well as the effects of modified gravity on them. We present the Fisher forecast formalism in Sec. V used to obtain the results in Sec. VI, which we summarize and discuss in Sec. VII.

II. HORNDESKI THEORY

We now describe the subclass of Horndeski theories studied in this paper. Horndeski theory [12] is the most general theory of gravity in four dimensions postulating a scalar field in addition to the metric tensor, while giving rise to second order equations of motion. Its Lagrangian

$$S[g_{\mu\nu},\phi]=\int \mathrm{d}^4x\sqrt{-g}\left[\sum_{i=2}^{5}\frac{1}{8\pi G_{\mathrm{N}}}\mathcal{L}_i[g_{\mu\nu},\phi]+\mathcal{L}_{\mathrm{m}}[g_{\mu\nu},\psi_M]\right] \tag{1}$$

contains four arbitrary functions $\{G_a(\phi,X), a=2,3,4,5\}$ for the scalar field ϕ,

$$\mathcal{L}_2 = G_2(\phi,X), \tag{2}$$

$$\mathcal{L}_3 = -G_3(\phi,X)\Box\phi, \tag{3}$$

$$\mathcal{L}_4 = G_4(\phi,X)R + G_{4X}(\phi,X)[(\Box\phi)^2 - \phi_{;\mu\nu}\phi^{;\mu\nu}], \tag{4}$$

$$\mathcal{L}_5 = G_5(\phi,X)G_{\mu\nu}\phi^{;\mu\nu} - \frac{1}{6}G_{5X}(\phi,X)[(\Box\phi)^3 + 2\phi_{;\mu}{}^{\nu}\phi_{;\nu}{}^{\alpha}\phi_{;\alpha}{}^{\mu} - 3\phi_{;\mu\nu}\phi^{;\mu\nu}\Box\phi], \tag{5}$$

where $X\equiv-\partial_\mu\phi\delta^\mu\phi$, $G_{\mu\nu}$ is the Einstein tensor, and R is the Ricci scalar. The matter Lagrangian is denoted by $\mathcal{L}_{\mathrm{m}}$ in Eq. (1), where $g_{\mu\nu}$ is the Jordan-frame metric, ψ_M are the matter fields, and G_N is the Newton gravitational constant. The partial derivatives are denoted with subscripts ϕ, X, e.g., $G_{5X}\equiv\partial G_5/\partial X$; the covariant derivatives are denoted with subscript ;.

An alternative and physically more meaningful basis of functions can be obtained by expanding the action to second order in linear perturbations of $g_{\mu\nu}$, ϕ, and other matter fields [13]. The action would consist of terms quadratic in the perturbations, each multiplied by time dependent functions which are only affected by the background cosmology, so that once the background expansion is fixed, the modifications to Einstein's equations in Horndeski theories are specified by four functions of time. A set of bases for these functions with direct physical interpretations was identified in Ref. [13] as α_i, $i\in\{M,T,B,K\}$:

(1) $\alpha_M\equiv \mathrm{d}\ln M_*^2/\mathrm{d}\ln a$, the running of the Planck mass, controls the strength of gravity given the initial value of M_*^2;
(2) $\alpha_T\equiv c_T^2-1$, the tensor speed excess, controls the excess speed of the gravitational waves propagation with respect to light;
(3) α_B, the braiding, parametrizes the mixing between the scalar field and metric kinetic terms [14];
(4) α_K, the kineticity, is the coefficient of the kinetic term for the scalar degrees of freedom before demixing [13].

The full form of the α's can be found in Appendix A. 3 of Ref. [15], and further explanations on these parameters can be found, e.g., in Ref. [13] and references therein.

Because of the gravitational wave observations of GW170817, α_T has been highly constrained [16,17], so we will fix it to practically zero throughout our work. We also cannot constrain α_K with subhorizon probes we are using in this work. Because α_K increases the relative strength of the kinetic to gradient term, it lowers the sound speed and hence the sound horizon to below the cosmological horizon, where a quasistatic configuration is reached [13,18], so α_K cannot be constrained with the quasistatic scales (although it can be probed with ultralarge scales [19]). As a result, we will focus solely on constraining α_M and α_B in this work.

We choose to restrict ourselves, following Ref. [20], to a class of models $\{\alpha_i\}$ whose time dependence follows the time evolution of the dark energy density:

$$\alpha_i(a) = c_i\Omega_{\mathrm{DE}}(a), \tag{6}$$

where a is the scale factor and c_i's are constants of proportionality. Note that this parametrization is purely phenomenologically motivated, and may imply fine-tuning between terms at the Lagrangian level.

We further restrict ourselves to consider matter that is minimally coupled to the metric without direct couplings to the scalar, so the effect of the modified gravity sector on matter is only mediated through the gravitational potential as in the case of general relativity. Finally, the background expansion is fixed to that of ΛCDM.

To compute the matter power spectrum of the considered models and the evolution of various quantities, we use the public code `hi_class`[5] [15,21]: Horndeski in CLASS (Cosmic Linear Anisotropy Solving System [22]). Our choice of parametrization corresponds exactly to the `propto_omega` option in `hi_class` that takes c_i's as input parameters, as well as the initial M_* which we set to 1 in units of the Planck mass $M_{\rm pl}$ to match GR solutions at early times.

III. POWER SPECTRUM OBSERVABLES

We now describe the power spectrum observables (galaxy clustering, lensing convergence, and their cross power) and how we obtain them given the linear matter power spectrum $P_m(k)$ from `hi_class`. We first define the observables in Sec. III A and describe their modeling in GR, and then we introduce the modifications from the Horndeski theory parameters in Sec. III B and study their impacts on the power spectra.

A. Definitions

A projected observable $X(\hat{\boldsymbol{n}})$ in full sky can be described in terms of its spherical harmonic transform X_l,

$$X(\hat{\boldsymbol{n}}) = \sum_{lm} X_{lm} Y_{lm}(\hat{\boldsymbol{n}}), \tag{7}$$

where $Y_{lm}(\hat{\boldsymbol{n}})$ denote the spherical harmonics. The angular power spectrum $C^{XZ}(l)$ is defined as

$$\langle X_{lm} Z^*_{l'm'} \rangle = \delta_{ll'}\delta_{mm'} C^{XZ}(l), \tag{8}$$

where $\langle \cdots \rangle$ denotes the ensemble average.

In the Limber approximation, the angular power spectrum between two probes X and Y is given by

$$C^{XZ}(l) = \int d\chi W^X(\chi) W^Z(\chi) \chi^{-2} P_m(k = l/\chi; \chi), \tag{9}$$

where $P_m(k;\chi)$ is the three-dimensional matter power spectrum, χ is the comoving distance, and W^X is the kernel for the probe X. In this work, we consider $X, Z \in \{\kappa, g\}$, where κ is lensing convergence and g is the galaxy density contrast.

Two galaxy samples are involved, the source sample which contains the background galaxies being lensed, and the lens sample which contains galaxies that act as a lens to the background galaxies. For our galaxy sample for g, we use the lens sample. We do not use a spectroscopic galaxy sample, though it is also possible to do so.

For galaxy density contrast g we have

$$W^g(\chi) \equiv b_1(\chi) \frac{p^{\rm len}(z(\chi))}{\bar{n}^{\rm len}} \frac{dz}{d\chi}, \tag{10}$$

where b_1 is the linear galaxy bias at χ, and $p^{\rm len}$ and $\bar{n}^{\rm len}$ are the redshift distributions and the average number density of the lens galaxy sample, respectively.

For the lensing convergence κ we have

$$W^\kappa(\chi) = \frac{3\Omega_{m,0} H_0^2}{2c^2 a(\chi) \bar{n}^{\rm src}} \int_\chi^{\chi_H} d\chi_s p^{\rm src}(z(\chi_s)) \frac{dz}{d\chi_s} \frac{\chi(\chi_s - \chi)}{\chi_s}, \tag{11}$$

where $p^{\rm src}$ and $\bar{n}^{\rm src}$ are the redshift distribution and the average number density of the source galaxy sample, respectively, $\Omega_{m,0}$ is the matter density at $z = 0$, H_0 is the Hubble constant today, and a is the scale factor.

We also consider tomography since the redshift dependence of the lensing kernel gives additional information that generally improves constraints. The power spectrum between X in redshift bin i and Z in redshift bin j is

$$C^{XZ}_{(ij)}(l) = \int d\chi W^X_{(i)}(\chi) W^Z_{(j)}(\chi) \chi^{-2} P_m(k = l/\chi; \chi), \tag{12}$$

where

$$W^g_{(i)}(\chi) \equiv b_{1,(i)}(\chi) \frac{p^{\rm lens}_{(i)}(z(\chi))}{\bar{n}^{\rm lens}_{(i)}} \frac{dz}{d\chi}, \tag{13}$$

where $b_{1,(i)}$ is the linear galaxy bias in the redshift bin i and

$$W^\kappa_{(i)}(\chi) = \frac{3\Omega_{m,0} H_0^2}{2c^2 a(\chi) \bar{n}^{\rm src}_{(i)}} \int_{\max\{\chi, \chi_i\}}^{\chi_{i+1}} d\chi_s p^{\rm src}_{(i)}(z(\chi_s)) \times \frac{dz}{d\chi_s} \frac{\chi(\chi_s - \chi)}{\chi_s}, \quad \text{for } \chi > \chi_{i+1}, \tag{14}$$

for $\chi \le \chi_{i+1}$ and

$$W^\kappa_{(i)}(\chi) = 0, \quad \text{for } \chi > \chi_{i+1}. \tag{15}$$

The quantities $p_{(i)}$ and $\bar{n}_{(i)}$ are now the redshift distribution and the average number density, respectively, in redshift bin i of the corresponding sample.

[5] `hi_class`: http://miguelzuma.github.io/hi_class_public/.

In principle, the galaxy bias is also a function of redshift and can be modeled, but we have chosen to model it as a nuisance parameter that varies with the redshift bin. So for three redshift bins, there are three values of $b_{1,(i)}$ to be marginalized over in the Fisher analysis, whereas for one redshift bin, there is only one. The fiducial values of $b_{1,(i)}$ are fixed at 1.

Note that for both the lens and source populations, we have ignored errors in the photometric redshifts for simplicity which is reasonable given that we consider only a few large redshift bins. We have also neglected the effects of intrinsic alignments, which would impact the constraints on cosmological parameters (see, e.g., Ref. [5]), with a degree that depends on how our understanding of systematics evolve in the next decade; we leave studying those effects to a future paper.

Now the observed power spectrum has additional noise contributions

$$\tilde{C}^{XZ}_{(ij)}(l) = C^{XZ}_{(ij)}(l) + N^{XZ}_{(ij)}, \tag{16}$$

where

$$N^{gg}_{(ij)} = \delta_{ij} \frac{1}{\bar{n}^{\rm lens}_{(i)}} \tag{17}$$

is the shot noise from the Poisson sampling of the underlying matter density for galaxies and

$$N^{\kappa\kappa}_{(ij)} = \delta_{ij} \frac{\sigma_\epsilon^2}{\bar{n}^{\rm src}_{(i)}} \tag{18}$$

accounts for the noise in the lensing convergence power spectrum due to the intrinsic ellipticity of the source galaxies. Furthermore, we assume $N^{g\kappa}_{(ij)} = 0$ here for simplicity.

For the particular Roman + LSST overlapping survey configuration, we start with the same distributions adopted in Ref. [5], which follows Ref. [7] in applying the exposure time calculator [23] on the CANDELS dataset—the detailed procedure may be found in Sec. 2.1 of Ref. [5] under item "Define the galaxy samples." Here we normalize the distributions to slightly different total number densities $\bar{n}^{\rm len} = \bar{n}^{\rm src} = 50$ arcmin^{-2} (cf. $\bar{n}^{\rm len} = 66$ arcmin^{-2} and $\bar{n}^{\rm src} = 51$ arcmin^{-2}), but we do not expect our results to be significantly impacted as they are dominated by the lensing convergence whose noise spectrum, controlled by $\bar{n}^{\rm src}$, is not significantly changed.

We clarify that this galaxy distribution results from an overlapping region of observation by Roman and LSST on the nominal Roman area of 2000 square degrees (this is strictly speaking, not a Roman + LSST forecast since we do not use all 18,000 square degrees of LSST). The reason for combining the two is that Roman is expected to observe a higher number density of sources at higher redshifts, hence yielding better shear measurements, while LSST provides better photometric redshift measurements. Note that we do not consider a LSST-only survey, which is also possible in principle, and would result in a wider but shallower survey. It is, however, beyond the scope of this paper to study the trade-offs here.

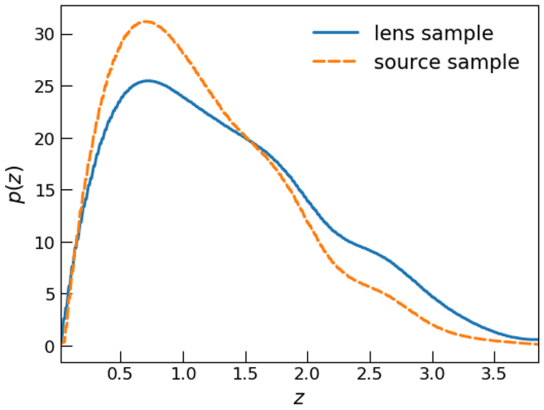

FIG. 1. Redshift distributions $p_{(i)}(z)$ for the overlapping 2000 deg^2 survey from Roman and LSST for the lens galaxy sample (blue solid curve) and source galaxy sample (orange dashed curve).

In Fig. 1 we show the redshift distributions $p(z)$ for the lens sample (blue solid curve) and the source samples (orange dashed curve) used in this paper over the notional 2000 deg^2 overlapping survey between the Roman and LSST. As stated previously, we do not include photometric redshift errors (contrary to Ref. [5]). Our boundaries for the redshift bins will be chosen such that the total number of lens galaxies inside each bin is the same, so that the galaxy noise spectra in Eq. (17) are constant between the redshift bins. As a result, the lensing convergence power spectrum will have slightly nonconstant noise spectra as the source sample distribution is slightly different from the lens sample distribution.

We model the covariance between the observed power spectra as

$$\begin{aligned}&{\rm Cov}[\tilde{C}^{XZ}_{(ij)}(l), \tilde{C}^{X'Z'}_{(ab)}(l')] \\ &= \frac{\delta_{ll'}}{(2l+1) f_{\rm sky}} [\tilde{C}^{XZ'}_{(ia)}(l)\tilde{C}^{ZZ'}_{(jb)}(l) + \tilde{C}^{XZ'}_{(ib)}(l)\tilde{C}^{ZX'}_{(ja)}(l)],\end{aligned} \tag{19}$$

where $i, j, a, b \in \{1 \cdots n_{\rm zbin}\}$ and $f_{\rm sky}$ is the fractional of the sky observed. For the overlapping 2000 deg^2 survey of Roman + LSST, we have $f_{\rm sky} = 0.0485$.

There is in principle a connected four-point function term due to the non-Gaussianity of the matter field, giving rise to the non-Gaussian covariance term, as well as a supersample covariance term due to the finite area of the survey. In Ref. [9], it was shown for the lensing power

spectrum (which dominates constraints in this study) that including non-Gaussian and supersample covariance could reduce the signal-to-noise ratio of the $C_l^{\kappa\kappa}$ by a factor of 2–3 in the $l_{\max} = 1000$–3000 range without tomography. However, when tomography is used, the reduction becomes less than a factor of 2. Furthermore, we recall that a factor of 2 in signal-to-noise corresponds to a much smaller change in the marginalized error of individual parameters. As shown in Ref. [24], a factor of at most 2 in signal-to-noise in the context of a eight-parameter Fisher analysis resulted in only a 10% change on the individual parameter constraints. This is because if the volume of the Fisher ellipsoid in higher dimensional space was to be shrunk by half and uniformly in all directions, then each of the N parameters would see their marginalized constraint change by a factor of $2^{1/N}$. So we will ignore non-Gaussian lensing contributions to the power spectrum covariance in this paper.

B. Power spectrum in modified gravity theories

The impact of modified gravity on the power spectrum observables can mainly be parametrized by two phenomenological effects μ and Σ in the quasistatic approximation, valid on scales much smaller than the cosmological horizon $H/k \ll 1$ and where the time derivatives of the perturbations are negligible compared to the spatial derivatives. In these limits, the quantity μ parametrizes the strength of the effective gravitational coupling $G_{\rm eff}$ in units of the Newton constant $G_N = 1/(8\pi M_{\rm pl}^2)$,

$$\mu = \frac{G_{\rm eff}}{G_N}, \tag{20}$$

which enters the modified Poisson equation that relates the gravitational potential Ψ to the matter density contrast δ_m,

$$\frac{k^2}{a^2}\Psi = -4\pi G\mu\rho_m\delta_m. \tag{21}$$

Consequently the matter power spectrum is modified with a growth function $D(k,\chi)$ that, unlike in GR, can now be in principle a scale-dependent function: $P_m(k;\chi) = D(k;\chi)^2 P_m(k)$.

Now the gravitational slip parameter γ is defined as

$$\gamma = \frac{\Phi}{\Psi}. \tag{22}$$

It follows from Eqs. (21) and (22) that

$$\frac{k^2}{a^2}(\Phi + \Psi) = 8\pi G\Sigma\rho_m\delta_m, \tag{23}$$

where

$$\Sigma = \frac{1}{2}\mu(1+\gamma). \tag{24}$$

The gravitational lensing is directly sensitive to Σ because it probes the combination $\Phi + \Psi$: the lensing kernels of Eq. (12) are modified as

$$W^{\kappa}_{(i)}(\chi) \to \Sigma(k = l/\chi;\chi) W^{\kappa}_{(i)}(\chi), \tag{25}$$

in the Limber approximation and could also inherit in principle a scale dependence through Σ. However, given the Horndeski theory adopted here with phenomenologically parametrized $\alpha_i(a) \propto \Omega_{\rm DE}(a)$, μ and Σ are only time dependent.

In the quasistatic limit with no anisotropic stress and assuming pressureless matter and negligible velocity perturbation on subhorizon scales, we can relate μ and Σ to α_i in Horndeski theories as follows[6] [25]:

$$\mu = \frac{M_{\rm pl}^2}{M_*^2}\frac{\alpha c_s^2(1+\alpha_T) + 2[-\alpha_B/(2(1+\alpha_T)) + \alpha_T - \alpha_M]^2}{\alpha c_s^2} \tag{26}$$

and

$$\gamma = \frac{\alpha c_s^2 - \alpha_B[-\alpha_B/(2(1+\alpha_T)) + \alpha_T - \alpha_M]}{\alpha c_s^2(1+\alpha_T) + 2[-\alpha_B/(2(1+\alpha_T)) + \alpha_T - \alpha_M]^2}, \tag{27}$$

where

$$\alpha = \alpha_K + \frac{3}{2}\alpha_B^2. \tag{28}$$

Note that α_K ends up dropping out of the expression for μ and γ as expected since its effects are not observable on quasistatic scales. We obtain the evolution of the quantities M_* and c_s from `hi_class` and compute μ and Σ using Eqs. (24), (26)–(28).

In the forecast work to follow, we actually fix the fiducial model to be not exactly but close to GR, to avoid the numerical singularity at $c_i = 0$. We adopt as fiducial MG parameters $\{c_B = c_M = c_T = 0.05, c_K = 0.1\}$. While nonzero α_M and α_T in the fiducial model means that a gravitational slip signal can be generated with a nonzero c_B (otherwise not present), we have verified that lowering the fiducial values to $\{c_M = c_B = 0.025, c_T = c_K = 0.005\}$ actually produces only slightly more constraining results. So the above choice is still a conservative one.

We show in Fig. 2 the impact on the lensing kernel W^{κ} by separately varying c_M (orange dashed curve) and

[6]See also "Notes on Horndeski Gravity" by Tessa Baker found at http://www.tessabaker.space.

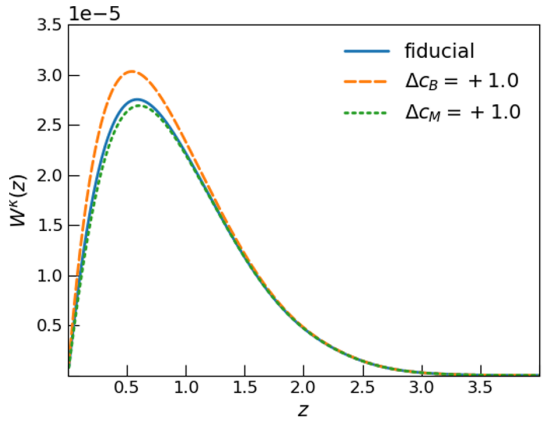

FIG. 2. Lensing kernel $W^\kappa(z)$ without tomography for the fiducial model (blue solid curve) and variations from it with $\Delta c_B = 1$ (orange dashed curve) and $\Delta c_M = 1$ (green dotted curve). The effects on the lensing kernel from c_B and c_M are opposite, resulting in opposite changes in the lensing-related power spectrum observables in Fig. 3.

c_B (green dotted curve) from the fiducial model (blue solid curve) by $\Delta c = +1.0$ in the case of no tomography. The differences are caused by the different time evolution of Σ in the two models, such that the lensing kernel is effectively weighted higher or lower.

In Fig. 3 we show the same effects on the power spectra. Note that the MG effects through the lensing kernel are bigger than that through the modified growth in the matter power spectrum, so we see again positive (negative) shifts for increased c_B (c_M) for the lensing-related power spectra $C^{\kappa\kappa}(l)$ and $C^{g\kappa}(l)$. On the one hand, C^{gg} is only sensitive to

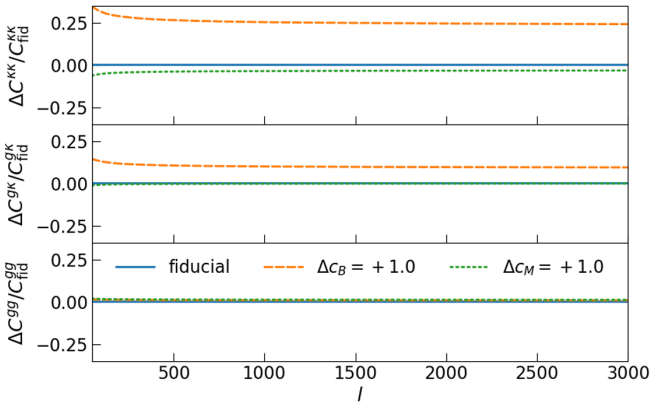

FIG. 3. Fractional deviation from the fiducial model for varying $\Delta c_B = 1$ (orange dashed curve) and $\Delta c_M = 1$ (green dotted curve) separately. The changes due to MG on the lensing probes $C^{\kappa\kappa}(l)$ and $C^{g\kappa}(l)$ are dominated by the effects through the lensing kernel rather than the matter power spectrum; on the other hand, the galaxy clustering probe $C^{gg}(l)$ is only sensitive to the much smaller effects of modified growth in the power spectrum. Furthermore, the fact that c_B and c_M induce opposite or same sign changes for the two kinds of power spectra means that combining them will help to break the degeneracy between c_B and c_M.

the matter power spectrum, on which the effects of c_B and c_M are much smaller and of the same sign. This would lead to negatively correlated c_B-c_M constraints with C^{gg} and positively correlated constraints with $C^{\kappa\kappa}(l)$ and $C^{g\kappa}(l)$. We expect therefore that combining all three probes could break parameter degeneracy in the c_B-c_M plane. Of course, after marginalizing over other cosmological parameters, the degeneracy directions shall become less sharply contrasted, but enough differences remain to yield improved constraints, as we shall see in Sec. VI.

IV. BISPECTRUM OBSERVABLES

A. Definitions

We follow Ref. [8] for the treatment of lensing convergence bispectrum and generalize to the case of cross-bispectra between any three observables

$$A, B, C \in \mathcal{O} \equiv \{\kappa_{(i)}, g_{(i)} | i = 1 \cdots n_{\rm zbin}\}, \tag{29}$$

where the fields κ and g in different redshift bins are treated as distinct observables.

For the bispectrum, we model the matter density fluctuation up to second order

$$\delta_m(\boldsymbol{k}) = \delta_m^{(1)}(\boldsymbol{k}) + \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \delta_m^{(1)}(\boldsymbol{q}) \delta_m^{(1)}(\boldsymbol{k}-\boldsymbol{q}) F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) + \mathcal{O}((\delta_m^{(1)})^3) \tag{30}$$

and the galaxies as a biased tracer of the matter field up to second order as well

$$\delta_g(\boldsymbol{x}) = b_1 \delta_m(\boldsymbol{x}) + \frac{b_2}{2} \delta_m^2(\boldsymbol{x}), \tag{31}$$

so that in Fourier space

$$\delta_g(\boldsymbol{k}) = b_1 \delta_m^{(1)}(\boldsymbol{k}) + b_1 \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \delta_m^{(1)}(\boldsymbol{q}) \delta_m^{(1)}(\boldsymbol{k}-\boldsymbol{q}) F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) + b_2 \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \delta_m^{(1)}(\boldsymbol{q}) \delta_m^{(1)}(\boldsymbol{k}-\boldsymbol{q}) + \mathcal{O}(\delta_m^{(3)}), \tag{32}$$

where

$$F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{5}{7} + \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_1 k_2} + \frac{2}{7} \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1 k_2} \tag{33}$$

is the second order perturbative kernel in GR.

The full-sky bispectra of the projected quantities $X, Y, Z \in \{\kappa, g\}$ in redshift bins $i, j, k \in \{1, \ldots, n_{\rm zbin}\}$, respectively, are defined as

$$\langle X_{l_1 m_1(i)} Y_{l_2 m_2(j)} Z_{l_3 m_3(k)} \rangle = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B^{XYZ}_{(ijk) l_1 l_2 l_3}, \tag{34}$$

where $\left(\begin{smallmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{smallmatrix}\right)$ is the Wigner-3j symbol that describes the coupling between the different modes. We approximate the Wigner-3j symbols by expanding the Stirling approximation to second order (see full expression in Appendix A), which is computationally fast and eliminates the accuracy problem for degenerate triangles in the commonly used first order expression.

The full-sky bispectrum is computed using the approximation that relates it to the flat-sky bispectrum,

$$B^{XYZ}_{(ijk) l_1 l_2 l_3} \approx \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \times B^{XYZ}_{(ijk)}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3), \tag{35}$$

where the flat-sky bispectrum is defined as

$$\langle X_{(i)}(\boldsymbol{l}_1) Y_{(j)}(\boldsymbol{l}_2) Z_{(k)}(\boldsymbol{l}_3) \rangle = (2\pi)^2 B^{XYZ}_{(ijk)}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3) \delta^D(\boldsymbol{l}_{123}), \tag{36}$$

where $\delta^D(\boldsymbol{l}_{123}) = \delta^D(\boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3)$. Here $X_i(\boldsymbol{l})$ is the Fourier transform of the projected observable in flat-sky,

$$X_{(i)}(\hat{\boldsymbol{n}}) = \sum_{\boldsymbol{l}} X_{(i),\boldsymbol{l}} e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}}, \tag{37}$$

where $\boldsymbol{l}$ is the Fourier wave vector.

In the presence of second order galaxy bias, the flat-sky bispectrum is composed of two pieces

$$B^{XYZ}_{(ijk)}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3) = \left[\int_0^{\chi_H} d\chi W^X_{(i)}(\chi) W^Y_{(j)}(\chi) W^Z_{(k)}(\chi) \chi^{-4} \times B_m(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3; \chi) \right] + B^{XYZ}_{(ijk),b_2}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3), \tag{38}$$

where $\boldsymbol{k}_i = \boldsymbol{l}_i/\chi$ in the Limber approximation. The first term is a projection of three-dimensional matter bispectrum at tree level where

$$B_m(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2P_m(k_1) P_m(k_2) F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) + 2\,\text{perm.}, \tag{39}$$

and where $F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is defined in Eq. (33) for GR and shall be modified for MG theories in Sec. IV B.

The second term comes from the second order galaxy bias b_2,

$$B^{XYZ}_{(ijk),b_2}(\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3) = \int_0^{\chi_H} d\chi \chi^{-4} \mathcal{I}^{XYZ}_{(ijk)b_2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3; \chi), \tag{40}$$

where

$$\mathcal{I}^{XYZ}_{b_2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = W^X_{(i),b_2} W^Y_{(j)} W^Z_{(k)} P(\boldsymbol{k}_2) P(\boldsymbol{k}_3) + W^X_{(i)} W^Y_{(j),b_2} W^Z_{(k)} P(\boldsymbol{k}_1) P(\boldsymbol{k}_3) + W^X_{(i)} W^Y_{(j)} W^Z_{(k),b_2} P(\boldsymbol{k}_1) P(\boldsymbol{k}_2). \tag{41}$$

The kernels $W^g_{(i)}$ and $W^\kappa_{(i)}$ involving b_1 are given by Eqs. (13) and (14), respectively, while those involving b_2 are given by

$$W^g_{(i),b_2} = b_{2,(i)} \frac{p^{\text{lens}}_{(i)}(z(\chi))}{\bar{n}^{\text{lens}}_{(i)}} \frac{dz}{d\chi} \tag{42}$$

and

$$W^\kappa_{(i),b_2} = 0. \tag{43}$$

As a result of Eq. (43), all $B^{\kappa\kappa\kappa}_{(ijk),b_2} = 0$. Furthermore, there is only one nonzero term for $\kappa_{(i)}\kappa_{(j)}g_{(k)}$ and its permutations, e.g.,

$$\mathcal{I}^{\kappa\kappa g}_{(ijk),b_2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = W^\kappa_{(i)} W^\kappa_{(j)} W^g_{(k),b_2} P(\boldsymbol{k}_1) P(\boldsymbol{k}_2), \tag{44}$$

two terms for $\kappa_{(i)}g_{(j)}g_{(k)}$ and its permutations, e.g.,

$$\mathcal{I}^{\kappa g g}_{(ijk),b_2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = W^\kappa_{(i)} \times (W^g_{(j),b_2} W^g_{(k)} P(\boldsymbol{k}_1) P(\boldsymbol{k}_3) + W^g_{(j)} W^g_{(k),b_2} P(\boldsymbol{k}_1) P(\boldsymbol{k}_2)), \tag{45}$$

and three for $g_{(i)}g_{(j)}g_{(k)}$.

Modeled as such, we have chosen to compute the lowest-order (nonloops) terms of the power spectra and the bispectra—P_{11} and the tree-level bispectrum, respectively. Note that b_2 did not appear in Sec. III A because it does not enter P_{11}. We have also chosen to ignore b_{s2} for simplicity.

The covariance between two general bispectra is given by Wick's theorem as

$$f_{\text{sky}} \text{Cov}\left[B^{XYZ}_{l_1 l_2 l_3 (ijk)}, B^{X'Y'Z'}_{l'_1 l'_2 l'_3 (abc)} \right] \approx \tilde{C}^{XX'}_{(ia)}(l_1) \delta_{l_1 l'_1} \left[\tilde{C}^{YY'}_{(jb)}(l_2) \tilde{C}^{ZZ'}_{(kc)}(l_3) \delta_{l_2 l'_2} \delta_{l_3 l'_3} + \tilde{C}^{YZ'}_{(jc)}(l_2) \tilde{C}^{ZY'}_{(kb)}(l_3) \delta_{l_2 l'_3} \delta_{l_3 l'_2} \right] + \tilde{C}^{XY'}_{(ib)}(l_1) \delta_{l_1 l'_2} \left[\tilde{C}^{YX'}_{(ja)}(l_2) \tilde{C}^{ZZ'}_{(kc)}(l_3) \delta_{l_2 l'_1} \delta_{l_3 l'_3} + \tilde{C}^{YZ'}_{(jc)}(l_2) \tilde{C}^{ZX'}_{(ka)}(l_3) \delta_{l_2 l'_3} \delta_{l_3 l'_1} \right] + \tilde{C}^{XZ'}_{(ic)}(l_1) \delta_{l_1 l'_3} \left[\tilde{C}^{YX'}_{(ja)}(l_2) \tilde{C}^{ZY'}_{(kb)}(l_3) \delta_{l_2 l'_1} \delta_{l_3 l'_2} + \tilde{C}^{YY'}_{(jb)}(l_2) \tilde{C}^{ZX'}_{(ka)}(l_3) \delta_{l_2 l'_2} \delta_{l_3 l'_1} \right], \tag{46}$$

where we have ignored non-Gaussian terms from connected 3-, 4-, and 6-point functions following Ref. [8] which verified that these terms are expected to be small over the angular range considered here for the lensing convergence bispectrum, which is the one that dominates our results as we shall see in Sec. VI.

The Kronecker delta functions in Eq. (46) enforce that the different triangles are uncorrelated. They also enforce that only one of the six terms is nonzero for a general triangle $l_1 \neq l_2 \neq l_3$, while two terms are present for isoceles triangles and all six for the equilateral triangles. Note that unlike for the auto-bispectrum, when considering the cross-bispectrum of different observables (e.g., $B^{\kappa gg}_{(123)}$), these two or six terms are not necessarily equal to each other anymore, so the full expression above shall be used, rather than the auto-bispectrum version:

$$f_{\rm sky}{\rm Cov}[B^{XXX}_{l_1l_2l_3(ijk)}, B^{XXX}_{l_1'l_2'l_3'(abc)}] \approx \Delta(l_1,l_2,l_3) C^{XX}_{(ia)}(l_1) C^{XX}_{(jb)}(l_2) C^{XX}_{(kc)}(l_3) \delta_{l_1l_1'} \delta_{l_2l_2'} \delta_{l_3l_3'}, \tag{47}$$

where $\Delta(l_1,l_2,l_3) = 1$, 2, or 6 for general, isoceles, and equilateral triangles.

In practice, we do not use all six terms but only keep the first of them for calculating the Fisher matrix. This is equivalent to treating all triangles as a general triangle $l_1 \neq l_2 \neq l_3$ even if they were actually equilateral or isoceles. The motivation behind this is that as we bin in l_1 and l_2 in the Fisher section, we expect that most of the triangles in a given l-bin represented by the bin center are not exactly equilateral or isoceles even if the triangle at the bin center happens to be one.

In principle, there are also additional non-Gaussian in-survey and supersample terms in the covariance between the bispectra as in the power spectrum case. In Ref. [9], it was shown that including these terms lead to at most a factor of 2–3 degradation in the signal-to-noise for combined lensing power spectrum and bispectrum for the range $l_{\rm max} = 1000$–3000, which further reduces to a factor of $\lesssim 2$ when tomography is used. A similar argument as the one made for the power spectrum constraints in Sec. III A applies here as well: the projected one-dimensional (1D) error on parameters would likely not exceed 15% when the higher-dimensional volume changes by a factor of 2–3 for a Fisher analysis with eight or more parameters.

B. Bispectrum in modified gravity theories

As described in Sec. III A, the effects of MG on the power spectrum observables mainly come through a modified growth of perturbations and gravitational slip which alters the matter power spectrum and the lensing kernel, respectively. These effects can be described by phenomenological parameters μ and Σ which are related to the α_i parameters in Horndeski theories. For the bispectrum, there is an additional effect through the second order perturbative kernel F_2 parametrized by λ [20]:

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = 1 + \frac{1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2} - \frac{2}{7}\lambda(a)\left(1-\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2}\right), \tag{48}$$

where $\lambda(a)$ obeys a second order differential equation [20,26,27]

$$\frac{{\rm d}^2\lambda}{{\rm d}\ln a^2} + \left(2+\frac{{\rm d}\ln H}{{\rm d}\ln a}+4f\right)\frac{{\rm d}\lambda}{{\rm d}\ln a} + (2f^2+\kappa_\Phi)\lambda = \frac{7}{2}(f^2-\tau_\Phi), \tag{49}$$

and $\lambda = 1$ in GR. Here H is the Hubble parameter, f is the linear growth rate whose evolution is also modified [20],

$$\frac{{\rm d}f}{{\rm d}\ln a} + \left(2+\frac{{\rm d}\ln H}{{\rm d}\ln a}\right)f + f^2 - \kappa_\Phi = 0, \tag{50}$$

and κ_Φ and τ_Φ encode the MG modifications to the usual equation describing gravitational evolution,

$$-\frac{k^2}{a^2H^2}\Phi(a,\boldsymbol{k}) = \kappa_\Phi(a,k)\delta(a,\boldsymbol{k}) + \int\frac{{\rm d}^3\boldsymbol{k}_1{\rm d}^3\boldsymbol{k}_2}{(2\pi)^3}\delta_{\rm D}^3(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{k})\tau_\Phi(a) \times\left(1-\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2}\right)\delta(a,\boldsymbol{k}_1)\delta(a,\boldsymbol{k}_2) + \mathcal{O}(\delta^3). \tag{51}$$

Instead of solving the differential equation for $\lambda(a)$, we follow Ref. [20] to use a phenomenological parametrization

$$\lambda(a) = \tilde{\Omega}_m^\xi(a), \tag{52}$$

where $\tilde{\Omega}_m(a)$ is the evolution of the matter density parameter and ξ is a fixed exponent. To leading order, ξ takes the following form:

$$\xi = \frac{-3+6\tilde{\gamma}+2\kappa_\Phi^{(1)}+7\tau_\Phi^{(1)}}{(7-6w^{(0)}+2c_{\rm M})(1-3w^{(0)}+c_{\rm M})}, \tag{53}$$

where $\tilde{\gamma}$ is the gravitational growth index $f \approx \tilde{\Omega}_m^{\tilde{\gamma}}(a)$ and $\tilde{\gamma} \approx 0.55$ in GR. The expressions for $\tilde{\gamma}$ in MG as well as for the lowest order expansion coefficients $\kappa_\Phi^{(1)}$, $\tau_\Phi^{(1)}$, $w^{(0)}$ as a function of c_i are given in Appendix B which are reproduced from Ref. [20].

C. Comments on the limitations of the adopted bispectrum modeling

There are a few limitations to the prescription used above to model MG effects on the bispectrum.

First, the prescription of modifying F_2 with $\lambda(a)$ is valid only for models in which the growth is a function of time alone [e.g., not valid for models such as $f(R)$ where the growth is also scale dependent]. In Ref. [18], the effects of screening on the power spectrum was modeled phenomenologically by introducing scale-dependent α's with a cutoff scale: $\alpha_i \to \alpha_i S(k/k_v)$ where $S(k/k_v) = \exp(-\frac{1}{2}(k/k_v)^2)$ where α's return to their GR values for scales smaller than the cutoff scale π/k_v. This would give the scale dependence that renders invalid the F_2 prescription adopted here for our bispectrum modeling.

Because any realistic MG models must pass the solar system tests with possibly a screening mechanism that returns the theory to GR on small scales, one might worry that not accounting for the screening would overestimate the amount of signal there is in reality on the small scales. It was, however, shown in Ref. [18] that introducing a screening scale as described above actually yields better constraints on parameters, as the existence of a new scale ends up contributing to break the degeneracies with other parameters. So although neglecting the screening effects here means not modeling the small scales accurately enough, it would actually lead to a more conservative, rather than optimistic forecast.

Second, the bispectrum modeling used here only includes the tree-level contribution, which is valid up to roughly $k \sim 0.1h^{-1}$ Mpc. For GR, a well-tested extension into the nonlinear regime exists where coefficients in front of the various terms in F_2 are added and fitted to GR simulations [28]. A similar extension into the nonlinear regime for MG is still being tested.

In Ref. [29], the authors combined the $\lambda(a)$ prescription together with the GR fitting formula to model the bispectrum in the nonlinear regime as

$$B^{\rm fit}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3;a) = 2P_m^{\rm NL}(k_1;a)P_m^{\rm NL}(k_2;a)F_2^{\rm fit}(\boldsymbol{k}_1,\boldsymbol{k}_2;a) + 2\ \text{perm.}, \tag{54}$$

where, compared to Eq. (39), the linear matter power spectrum has now been replaced by the nonlinear power spectrum $P_m^{\rm NL}$ in MG, and where the F_2 kernel now includes nonlinear effects through the coefficients $\bar{a}, \bar{b}$, and $\bar{c}$ which are fitted on GR simulations:

$$F_2^{\rm fit}(\boldsymbol{k}_1,\boldsymbol{k}_2;a) = \left(1-\frac{2}{7}\lambda(a)\right)\bar{a}(k_1,a)\bar{a}(k_2,a) + \frac{1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\bar{b}(k_1,a)\bar{b}(k_2,a) + \frac{2}{7}\lambda(a)\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2}\bar{c}(k_1,a)\bar{c}(k_2,a). \tag{55}$$

The validity of this formula is tested against simulations in Ref. [29] for the $f(R)$ and the Dvali–Gabadadze–Porrati models in the equilateral triangle configurations. More validation work is to be done for other MG models as well as for general triangle configurations. While this work is in progress, we restrict ourselves to the modeling at the tree level which becomes less valid in the nonlinear regime. We will control the degree to which this affects our results by varying the angular scale cuts $l_{\rm max}$ of our Fisher results in Sec. VI, and note that we expect better constraints once the nonlinear regime can be properly modeled.

V. FORECAST SETUP

We now describe the Fisher matrix formalism used to obtain the results later presented in Sec. VI. We first set up the set of observables to be used in Sec. V A and describe the Fisher matrix formulas in Sec. V B.

A. The complete set of unique bispectra

One needs to be careful when finding the unique and complete set of cross-bispectrum observables when using tomography. The κ and g fields in each redshift bin are now counted as a unique observable. So for $n_{\rm zbin} = 3$ redshift bins and 2 fields, we really have 6 different observables $\mathcal{O} = \{\kappa_{(i)}, g_{(i)} | i = 1..n_{\rm zbin}\}$. The bispectra between them can then be separated into three categories:

(1) Pure auto-bispectra, such as $\kappa_{(1)}\kappa_{(1)}\kappa_{(1)}$ or $g_{(2)}g_{(2)}g_{(2)}$;
(2) Cross-bispectra where two of the observables are the same, such as $\kappa_{(1)}g_{(2)}g_{(2)}$ or $\kappa_{(3)}\kappa_{(3)}g_{(1)}$;
(3) Cross-bispectra between three completely different observables, such as $\kappa_{(1)}\kappa_{(2)}g_{(3)}$.

Note that if we counted only unique triangles $\{(l_1,l_2,l_3)|l_1 \le l_2 \le l_3\}$ for each combination ABC where $A,B,C \in \mathcal{O}$, then in case 1 ($A = B = C$), any of the six possible permutations of ABC would be redundant. However, in case 3 (where all three observables A, B, and C are distinct), then all six permutations of ABC are unique. Finally, we have case 2 which are the intermediate cases ($A = B$, $A = C$, or $B = C$), where the three *cyclic* permutations of ABC form a unique set. To account for all of this, we adopt only the unique permutations of ABC for each case as described above. It would be equivalent to permute (l_1,l_2,l_3) instead of ABC, but because this would result in different sets of triangles to be looped over for each kind of bispectrum, we find it easier in practice to not do so. We will implement the loop over multipoles as an "outer loop."

In general, there are a total of $n_{\rm obs}^3$ distinct bispectra; for $n_{\rm zbin} = 3$ this would be 216. We can, however, reduce this number in our case by noticing that $g_{(i)}g_{(j)}g_{(k)}$ is nonzero only if we are considering the same redshift bin $i = j = k$. Because the lensing kernels are nonzero over the redshift range up to the bin considered, a further reduction can be done by keeping only the bispectra in which κ are not from a redshift bin lower than the lowest bin for any g.

This reduces the total number of bispectra to model to 90 for $n_{\rm zbin}=3$ and 34 for $n_{\rm zbin}=2$.

We note also that oftentimes in the literature, a redefinition of the bispectrum in case 2 is used, when the bispectrum is invariant under cyclic permutations of ABC. For example, for

$$B^{\kappa\kappa g}_{(123)}(l_1,l_2,l_3)=B^{g\kappa\kappa}_{(312)}(l_1,l_2,l_3)=B^{\kappa g\kappa}_{(231)}(l_1,l_2,l_3), \quad (56)$$

one would possibly redefine $B^{\kappa\kappa g}_{(123)}(l_1,l_2,l_3)$ to mean just the sum of all three bispectra,

$$B^{\kappa\kappa g}_{(123)}(l_1,l_2,l_3)\to 3B^{\kappa\kappa g}_{(123)}(l_1,l_2,l_3), \quad (57)$$

and deal with fewer bispectrum observables. In this work, however, the invariance is broken by the presence of the second order bias b_2 terms.[7] But even if we were to not include those b_2 terms, because the covariance with other bispectra is not invariant under cyclic permutations (need to match l_1 with l_1, etc.), we would still need to spell out the individual bispectrum in the definition. For these reasons, we adopt the conventions described above.

In Figs. 4–6, we show the impact of varying the parameters c_M, c_B, and b_2 on all eight bispectra found in the case of $n_{\rm zbin}=1$. In the top panel, we show the flat sky bispectrum signal in the fiducial model as a function of triangle configurations for $\kappa_{(1)}\kappa_{(1)}\kappa_{(1)}$ (blue solid lines) and $g_{(1)}g_{(1)}g_{(1)}$ (orange dashed lines). We do not show the other bispectra as they all have similar shapes with amplitudes between the two shown curves.

The triangle configurations are ordered by increasing l_1, l_2, and then l_3. The gray vertical lines denote where l_1 steps up; the red squares where l_2 steps up or down also correspond to the isoceles triangles with $l_3=l_2$ or near-isoceles triangles with $l_3=l_2+1$ (for when the bispectra with $l_3=l_2$ is zero due to Wigner-3j symbols vanishing for $l_1+l_2+l_3=$ odd); finally the green dots denote $l_2=l_1$ isoceles triangles.

We show in the lower panels the fractional deviation from the fiducial bispectrum signal. For c_M and c_B, all three cyclic permutations of case 2 ABC (panels 4–6 and 7–9) have the same fractional deviations; this is not true for the $b_{2,(1)}$ parameter as mentioned above.

B. The Fisher matrix formalism

Given a set of bispectra $\mathcal{M}$, the Fisher matrix from combining all of them is given by

$$F^{\rm B}_{\alpha\beta}=\sum_{l_{\rm min}\le l_1\le l_2\le l_3\le l_{\rm max}}\sum_{M,M'\in\mathcal{M}}\frac{\partial B^M_{l_1l_2l_3}}{\partial p_\alpha}\left[{\rm Cov}\left[B^M_{l_1l_2l_3},B^{M'}_{l_1l_2l_3}\right]\right]^{-1}\times\frac{\partial B^{M'}_{l_1l_2l_3}}{\partial p_\beta}, \quad (58)$$

where we considered the parameters p_α and p_β and ${\rm Cov}[X,Y]$ denote the covariance matrix defined in Eq. (46). In this work, we will obtain results for the entire set of unique and nonzero bispectra described in the previous section $\mathcal{M}=\mathcal{M}^{\rm tot}$, as well as for various subsets of $\mathcal{M}^{\rm tot}$ such as those containing only κ, only g, or only auto-bispectra. Note that because of the use of tomography, an auto-bispectrum is no longer anything of the form $\kappa\kappa\kappa$ or ggg, but rather $\kappa_{(i)}\kappa_{(i)}\kappa_{(i)}$ or $g_{(i)}g_{(i)}g_{(i)}$ with fields belonging to the same redshift bin.

For our fiducial results, we will take $n_{\rm zbin}=3$ for which the covariance matrix is a 90×90 matrix for each triplet (l_1,l_2,l_3). Recall that the bispectrum covariance is diagonal in triangle configuration space, so we only need to sum over pairs of the same triangle configuration (l_1,l_2,l_3) and make sure to count each configuration only once by imposing $l_1\le l_2\le l_3$.

Since it is intractable to compute contributions from every l, we follow Ref. [8] to bin l_1 and l_2,

$$F^{\rm B}_{\alpha\beta}=\sum_{l_{\rm min}\le\bar{l}_1\le\bar{l}_2\le l_3\le l_{\rm max}}\Delta l_1\Delta l_2\times\sum_{M,M'\in\mathcal{M}}\frac{\partial B^M_{\bar{l}_1\bar{l}_2l_3}}{\partial p_\alpha}\left[{\rm Cov}\left[B^M_{\bar{l}_1\bar{l}_2l_3},B^{M'}_{\bar{l}_1\bar{l}_2l_3}\right]\right]^{-1}\frac{\partial B^{M'}_{\bar{l}_1\bar{l}_2l_3}}{\partial p_\beta}, \quad (59)$$

where $\bar{l}_1$ and $\bar{l}_2$ denote the logarithmic center of the logarithmic bins in l_1 and l_2. As pointed out by Ref. [8], since the Wigner-3j symbol is only nonzero for $\bar{l}_1+\bar{l}_2+l_3=$ even and vanishing for $\bar{l}_1+\bar{l}_2+l_3=$ odd, and the nonzero values themselves change rapidly in sign when varying l_3 with fixed l_1 and l_2, one should not bin over l_3 for accurate results.

Note that for the purpose of calculating the Wigner-3j symbols, $\bar{l}_1$ and $\bar{l}_2$ are actually the nearest integers to the log bin centers, which is a good enough approximation if the bins are large enough to cover multiple integers. For all our results, we use 26 logarithmic l bins between $l_{\rm min}=50$ and $l_{\rm max}=3000$. This is reasonable as the bispectra considered here are smooth enough between the bins chosen.

For the power spectrum Fisher matrix, we use

$$F^{\rm PS}_{\alpha\beta}=\sum_{\bar{l}}\Delta l\sum_{N,N'\in\mathcal{N}}\frac{\partial C^N_{\bar{l}}}{\partial p_\alpha}\left[{\rm Cov}\left[C^N_{\bar{l}},C^{N'}_{\bar{l}}\right]\right]^{-1}\frac{\partial C^{N'}_{\bar{l}}}{\partial p_\beta}, \quad (60)$$

where $\mathcal{N}=\{\kappa_{(i)}\kappa_{(j)},g_{(i)}\kappa_{(j)},g_{(i)}g_{(j)}|i,j=1,2,\ldots,n_{\rm zbin}$ and $i\le j\}$ is a set of unique auto- and cross-power spectra.

[7]Note that the sum of those three terms in Eq. (56) can still be expressed as invariant under cyclic permutations as long as b_2 and b_1 do not evolve with redshift. Here this form is explicitly broken as we model b_2 and b_1 to change between tomographic bins.

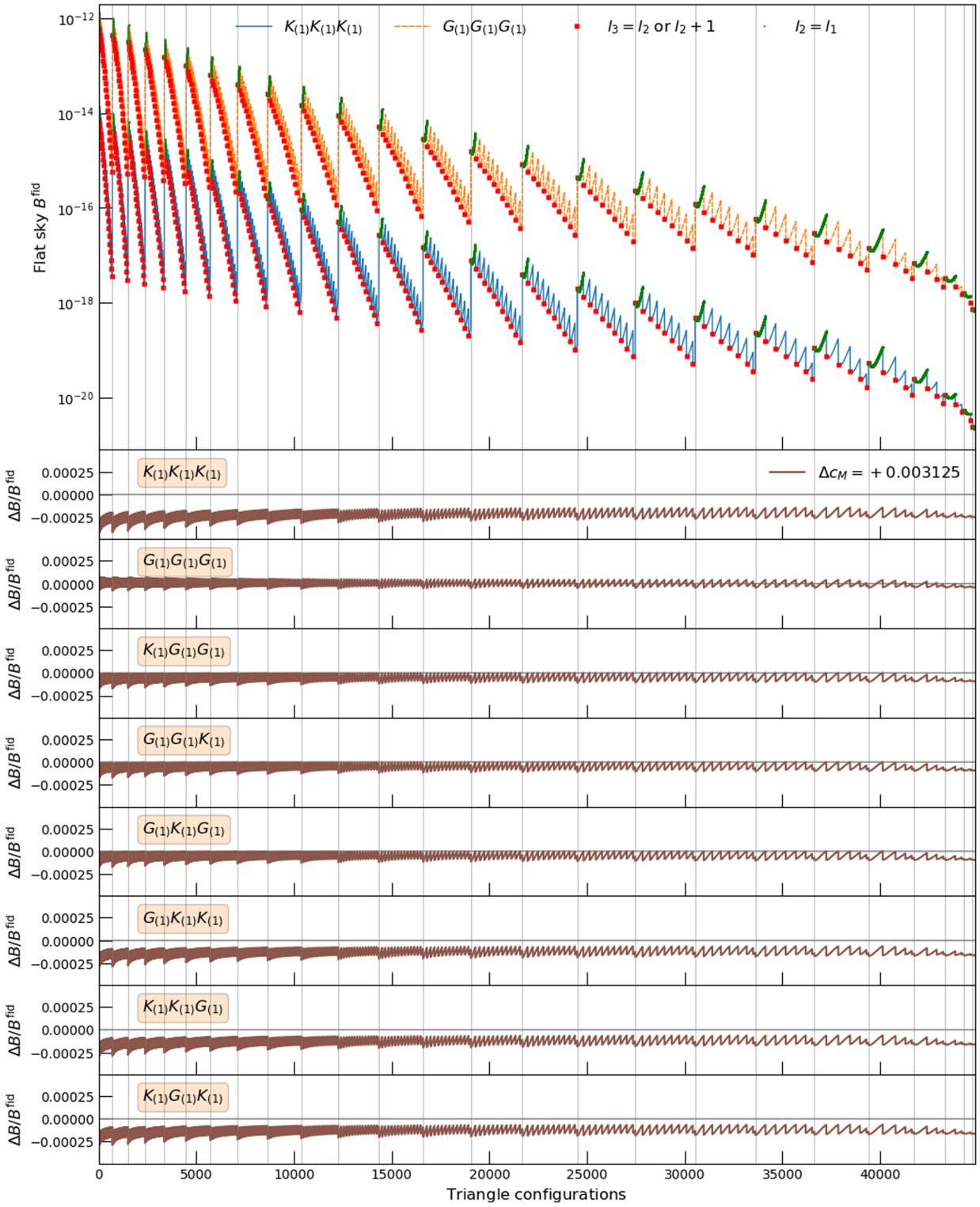

FIG. 4. Top panel: Flat sky bispectrum signal in the fiducial model vs triangle configurations for two out of eight bispectra combinations for $n_{\rm zbin} = 1$: $\kappa_{(1)}\kappa_{(1)}\kappa_{(1)}$ (blue solid line) and $g_{(1)}g_{(1)}g_{(1)}$ (orange dashed line); all other bispectra (not shown) have similar shapes but with amplitudes somewhere between these two. We order the triangle configurations first by increasing l_1 (l_1 steps up at gray vertical lines), then l_2 (which steps up or down at red squares), and then l_3. The red squares also mark the isoceles triangles $l_3 = l_2$ or near-isoceles triangles $l_3 = l_2 + 1$ (when the bispectra with $l_3 = l_2$ is zero due to vanishing Wigner-3j symbols for $l_1 + l_2 + l_3 =$ odd); the green dots mark the $l_2 = l_1$ isoceles triangles, and for a given l_1, they form a line along which l_3 is increased. Lower panels: Fractional deviations from the fiducial bispectrum signal for all eight bispectra in the $n_{\rm zbin} = 1$ case, when c_M is varied by $\Delta c_M = 0.003125$ (as used in the derivative computation).

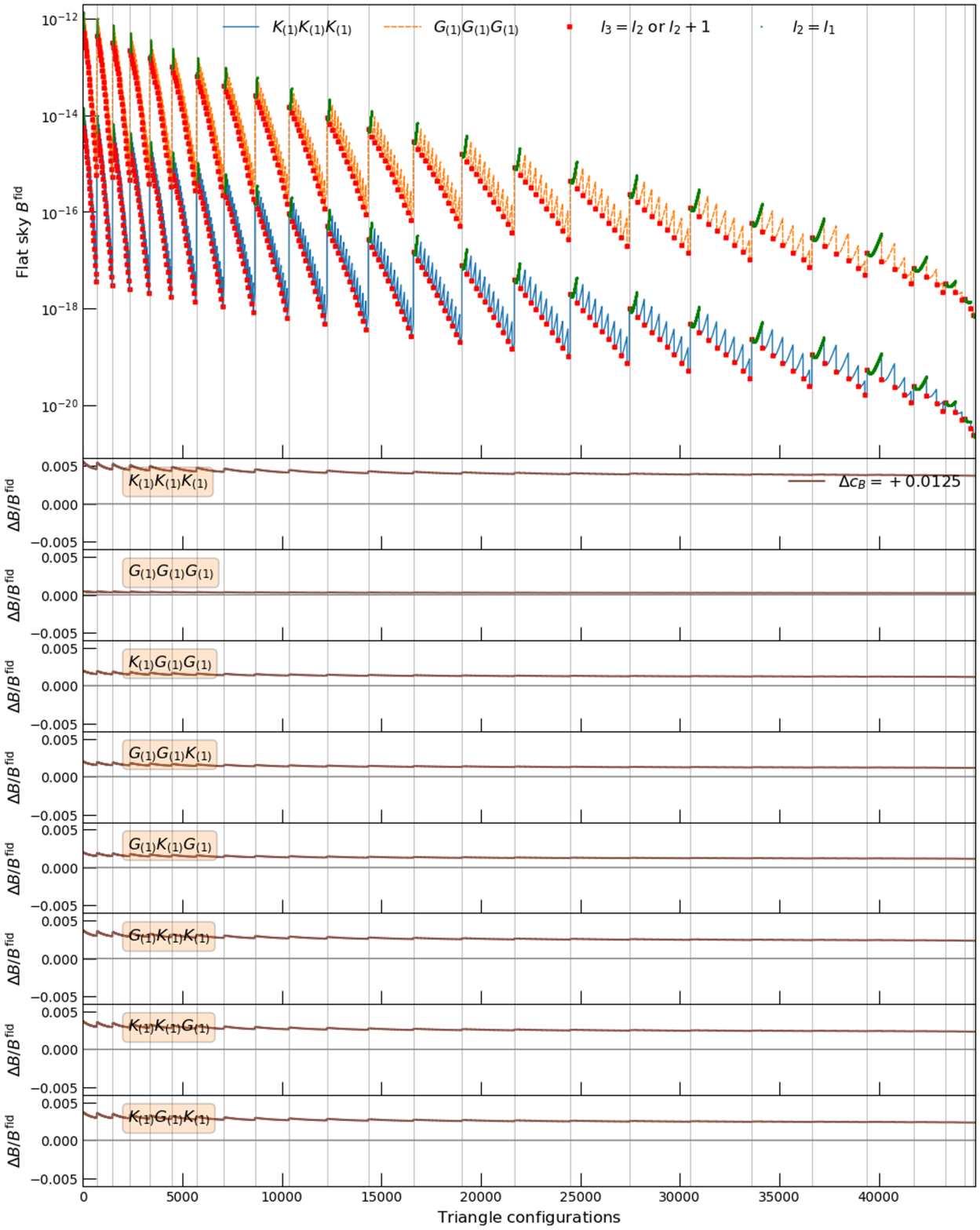

FIG. 5. Same as Fig. 4, but for c_B varying by $\Delta c_B = 0.0125$.

Note that we have ignored cross-spectra of the form $g_{(i)}\kappa_{(j)}$ where $i > j$, i.e., where the κ bin is in front of the galaxy bin. These spectra would be nonzero in reality in the presence of magnification effects and photometric redshift outliers, which we consider as systematics worth investigating in the future and omit them here for simplicity. Therefore, the covariance matrix at a given $\bar{l}$ is a $|\mathcal{N}| \times |\mathcal{N}|$ matrix given by Eq. (19) where $\mathcal{N} = 3n_{\rm zbin}(n_{\rm bin} + 1)/2$.

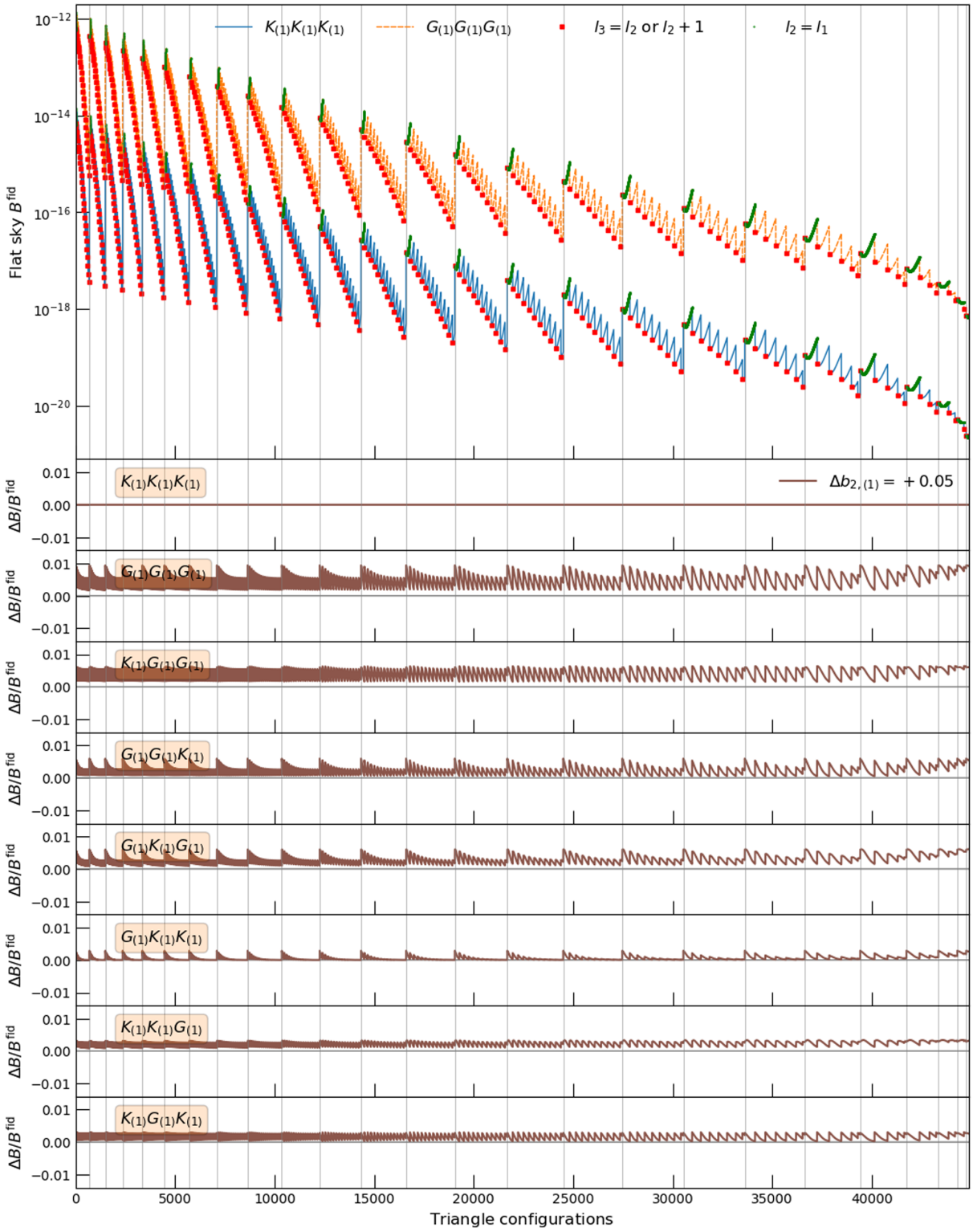

FIG. 6. Same as Fig. 4, but for $b_{2,(1)}$ varying by $\Delta b_{2,(1)} = 0.05$. Note that while the fractional deviations are the same for cyclically permuted bispectra (panels 4–6 and 7–9) for c_M and c_B, it is not so when $b_{2,(1)}$ is varied.

To compute the total Fisher information combining the power spectra and bispectra, we simply add the two Fisher matrices and ignore correlations between them

$$F^{\rm tot}_{\alpha\beta} \approx F^{\rm PS}_{\alpha\beta} + F^{\rm B}_{\alpha\beta}. \tag{61}$$

In principle, there are additional correlations arising from the 5-point function between the observables, which could degrade the total constraints. We leave its consideration for future work and focus on our aim of estimating the relevance of cross-bispectra for this work.

Under the approximation that the likelihood function is a multivariate Gaussian, the inverse of the Fisher matrix gives us the covariance between any two measured parameters α and β,

$$\mathrm{Cov}[\alpha,\beta] = \begin{bmatrix} (\boldsymbol{F}^{-1})_{\alpha\alpha} & (\boldsymbol{F}^{-1})_{\alpha\beta} \\ (\boldsymbol{F}^{-1})_{\alpha\beta} & (\boldsymbol{F}^{-1})_{\beta\beta} \end{bmatrix}. \tag{62}$$

We use this covariance matrix to plot the two-dimensional (2D) contours on parameter constraints in Sec. VI. Moreover, the 1D marginalized constraints on a parameter α will be given by

$$\sigma_\alpha = [(\boldsymbol{F}^{-1})_{\alpha\alpha}]^{1/2}. \tag{63}$$

Throughout this work, we use a ΛCDM model consistent with the final Planck 2018 results (baseline model 2.5): primordial spectrum amplitude and tilt $A_s = 2.1 \times 10^{-9}$ and $n_s = 0.966$ (for the pivot scale $k_* = 0.05\ \mathrm{Mpc}^{-1}$), Hubble constant $h_0 = 0.673$, matter density $\Omega_m = 0.316$, and baryon density $\Omega_b = 0.0494$. The derivatives are computed around a modified gravity fiducial model very close to GR as mentioned before: $c_K = 0.1$, $c_B = 0.05$, $c_M = 0.05$, $c_T = 0.05$. This is chosen so that we can avoid numerical singularities at $c_i = 0'$.

The derivatives are calculated using a two-sided finite difference by varying the following set of parameters one at a time from their fiducial values (c_T and c_K are always fixed): $\{c_M, c_B, \Omega_m, b_{j,(i)}, A_s, \Omega_b, h_0, n_s | i = 1, \ldots, n_{\rm zbin}\}$ where we have only linear galaxy biases $j = 1$ for the power spectra, and both linear and second order biases $j = 1, 2$ for the bispectra. The fiducial values for the linear galaxy biases are fixed at $b_{1,(i)} = 1$ for all redshift bins, while the fiducial second order biases are computed using a fitting formula $b_2(b_1)$ derived from GR simulations [Eq. (5.2) of Ref. [30]] by evaluating it at the fiducial value $b_1 = 1$; however, we do not vary b_2 with this formula when we vary b_1 for the derivative computation so to treat them as separately measured parameters.

Now most of the parameters are varied with a step size 5% of their fiducial value, except for c_M, c_B, and Ω_m where we used $\Delta c_M = 0.003125$, $\Delta c_B = 0.0125$, and $\Delta\Omega_m = 0.004$ to guarantee convergence of the derivatives. For the four parameters at the end of the list, we imposed priors consistent with the Planck 2018 constraints: they are $\sigma_{A_s} = 3.1 \times 10^{-11}$, $\sigma_{\Omega_b} = 3.3 \times 10^{-4}$, $\sigma_{h_0} = 0.006$, and $\sigma_{n_s} = 0.0044$. These priors are included by adding a diagonal matrix $(F_{\rm prior})_{\alpha\beta} = \delta_{\alpha\beta}\sigma_\alpha^{-2}$.

VI. RESULTS

We will now show the results of the Fisher forecasts. We will first study the constraints from the bispectra alone in Sec. VI A, by breaking down the results from all bispectra into those from smaller subsets, showing how using all cross-bispectra between κ and g in various redshift bins helps to improve parameter constraints. Then we will look at the total results from combining the power spectra and bispectra in Sec. VI B, as well as the dependence on some forecast parameters: the number of redshift bins and $l_{\rm max}$. Unless otherwise mentioned, we report our results at a fiducial choice of $l_{\rm max} = 1000$ and $n_{\rm zbin} = 3$.

A. Bispectrum results

The marginalized 2D parameter constraints for c_M, c_B, and Ω_m at 68% confidence level are shown in Fig. 7 for the set of all unique and nonzero bispectra as well as for a few informative subsets.

Restricting $\mathcal{M}$ in Eq. (59) to $\mathcal{M} = \{M \in \mathcal{M}^{\rm tot} | M = \kappa_{(i)}\kappa_{(j)}\kappa_{(k)}\}$ give the $\kappa\kappa\kappa$ only results, and similarly for

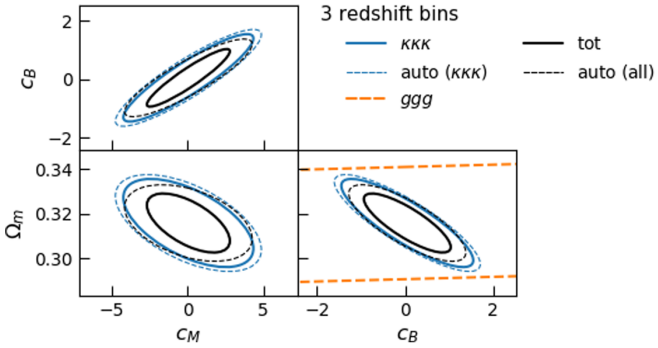

FIG. 7. Marginalized 2D parameter constraints shown at 68% confidence level from all bispectra and chosen subsets. Results are shown for our fiducial choice of $l_{\rm max} = 1000$ and $n_{\rm zbin} = 3$, and the 2000 deg^2 area of overlap between the notional Roman Space Telescope survey and LSST. The $\kappa\kappa\kappa$ constraint (blue solid curve) does much better than the ggg constraint (orange dashed curve) which is too big to show for some of the panels. Combining them (black solid curve) results in a significant improvement, a factor of (8, 52, 3) for the constraints on (c_M, c_B, Ω_m) compared to ggg alone and (1.6, 1.5, 1.5) for $\kappa\kappa\kappa$ alone. Furthermore, while the $\kappa\kappa\kappa$ result is dominated by the κ auto-bispectra, the total bispectrum result is not dominated by the auto-bispectra alone (black dashed curve), indicating that including cross-bispectra between $\kappa(i)$ and $g(j)$ serves to improve constraints, quantitatively, by a factor of (1.3, 1.1, 1.3) on c_M, c_B, and Ω_m constraints, respectively.

TABLE I. Marginalized 1σ parameter constraints from power spectra (P), bispectra (B), and combined power spectra and bispectra (P + B) for a notional survey of 2000 deg^2 of the Roman Space Telescope survey overlapped with LSST. We included Planck 2018 priors on the parameters $\{A_s, n_s, \Omega_b h^2, h\}$. The second order biases $b_{2,(i)}$ (not shown here) are also marginalized over for the bispectrum constraints. Improvements of the combined results over power spectra alone are a factor of ∼1.3 for $n_{\rm zbin} = 1$, and a factor of ∼1.2 for $n_{\rm zbin} = 2$ and 3 in both parameters c_M and c_B.

	No tomography			2 Redshift bins			3 Redshift bins		
	P	B	P + B	P	B	P + B	P	B	P + B
$\sigma(c_M)$	2.3	3.0	1.8	1.4	2.3	1.2	1.2	1.8	0.98
$\sigma(c_B)$	0.79	1.1	0.63	0.46	0.85	0.40	0.37	0.65	0.31
$\sigma(\Omega_m)$	0.009	0.016	0.008	0.006	0.011	0.006	0.005	0.009	0.005
$b_{1,(1)}$	0.017	0.045	0.015	0.009	0.024	0.008	0.007	0.020	0.007
$b_{1,(2)}$	⋯	⋯	⋯	0.011	0.051	0.010	0.007	0.025	0.007
$b_{1,(3)}$	⋯	⋯	⋯	⋯	⋯	⋯	0.010	0.054	0.009

ggg alone. They correspond to the solid blue line and the dashed orange line, respectively. The *ggg* only contours are too big to show in some of the panels—about ∼5 times bigger in the c_M direction and about ∼35 times bigger for c_B. The reason c_B does much worse with *ggg* is because the lensing kernel is much more sensitive to c_B than the growth of perturbations, as was also the case for the power spectrum shown in Fig. 3. This is also evident from Fig. 5, when comparing for example the fractional deviation curves for the $\kappa_{(1)}\kappa_{(1)}\kappa_{(1)}$ and $g_{(1)}g_{(1)}g_{(1)}$ bispectra in the second and third panels.

The total result combining both kinds of probes (black solid line) gives an improvement a factor of (1.6, 1.5, 1.5) over the *κκκ* only results for the (c_M, c_B, Ω_m) constraints, and a factor of (8, 52, 3) over *ggg* alone for the same parameters. This provides motivation for combining both

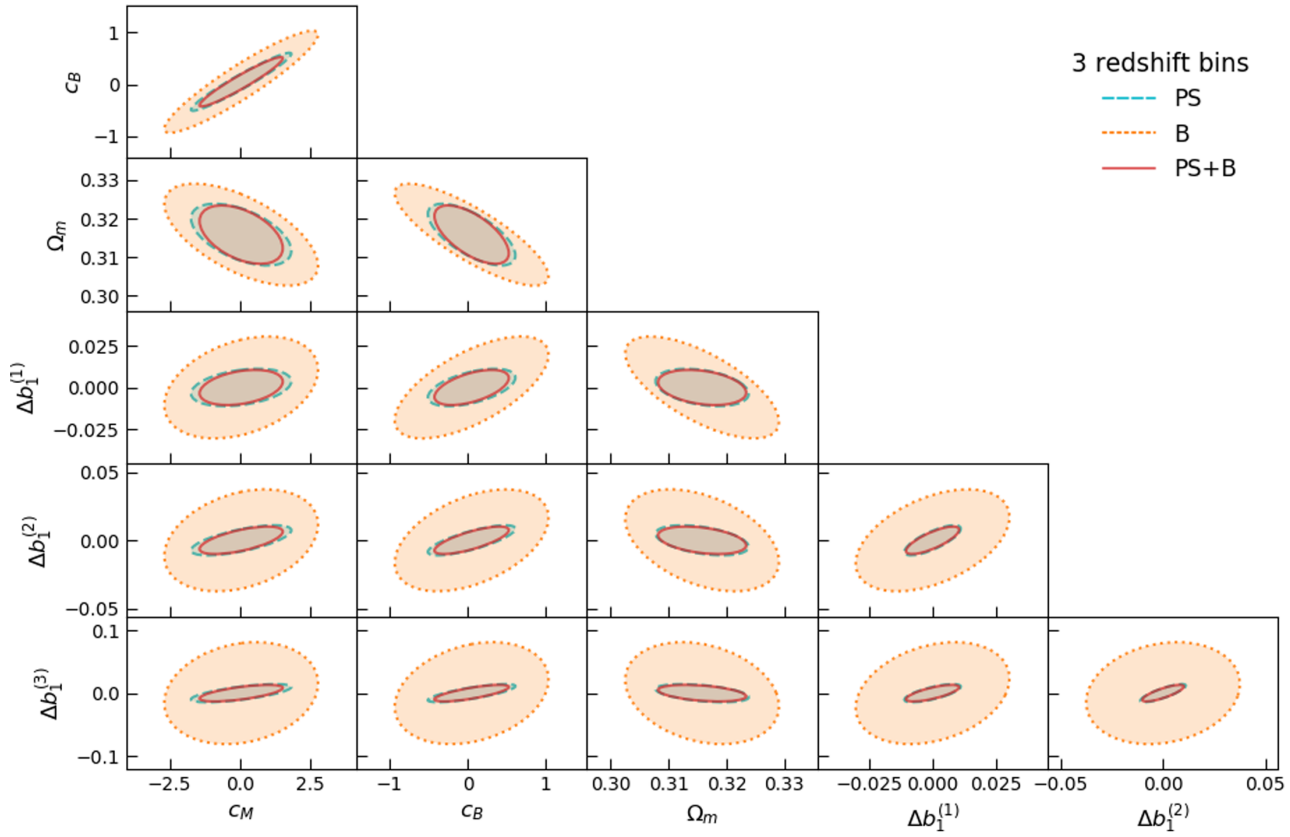

FIG. 8. Marginalized 2D parameter constraints at 68% confidence level from the power spectra alone (PS, green dashed curve), the bispectra alone (B, orange dotted curve), and their combination (PS + B, red solid curve) for the fiducial $l_{\rm max} = 1000$ and $n_{\rm zbin} = 3$ setup. The combined PS + B constraints improve on power spectra alone by about a factor of ∼1.2 for both c_M and c_B.

the lensing convergence and galaxy density probes in a bispectrum analysis for Horndeski models.

Furthermore, we see that the use of cross-bispectra is important for obtaining such results. While the $\kappa\kappa\kappa$ constraints are dominated by its subset of auto-bispectra (dashed blue line), the total result is not dominated by simply combining all the auto-bispectra (dashed black line) where $\mathcal{M} = \{M = \kappa_{(i)}\kappa_{(i)}\kappa_{(i)} \text{ or } g_{(i)}g_{(i)}g_{(i)} | M \in \mathcal{M}^{\text{tot}}\}$. (We did not show auto-bispectra for ggg, since this is already a set of auto-bispectra by definition; see Sec. VA.) It is the inclusion of all the cross-bispectra between $\kappa(i)$ and $g(j)$ that contribute to improving constraints over auto-bispectra alone—a factor of (1.3, 1.1, 1.3) improvement on the (c_M, c_B, Ω_m) errors, respectively.

We also list in Table I the 1D marginalized constraints at 68% confidence level for the total bispectrum results for various parameters. In particular, we have $\sigma_{c_M} = 1.8$ and $\sigma_{c_B} = 0.65$ for $n_{\text{zbin}} = 3$.

B. Combined power spectrum and bispectrum results

We now proceed to combine the power spectrum and bispectrum results. We show in Fig. 8 the power spectrum (PS) constraints in green dashed curves, the bispectrum (B) in orange dotted curves, and the combined PS + B in red solid curves, here again for our fiducial choice of $l_{\text{max}} = 1000$ and $n_{\text{zbin}} = 3$. The bispectrum does worse on its own than the power spectrum constraints alone, but adding the bispectra improves the constraints on both MG parameters by a factor of ∼1.2 compared to PS alone, with modest improvement (about 1.1) for the other parameters without priors ($\{\Omega_m, b_{1,(i)}\}$). The same kind of improvement is observed for $n_{\text{zbin}} = 2$ (not shown here).

The degeneracy directions in the c_M-c_B plane for the bispectra and power spectra are surprisingly similar. They are, however, more visibly different for the bias parameters. This makes sense since the power spectra and bispectra are proportional to different powers of the galaxy bias. It seems that the improvement in MG parameters mostly comes from the breaking of degeneracy in the bias planes, and that better bispectrum constraints in those planes would lead to improved MG parameter constraints as well.

In fact, we see in Fig. 9 where we let l_{max} increase from 1000 to 3000 that the bispectrum constraints are closer to those of the power spectrum, because of the much larger number of triangles available at higher multipoles compared to the PS modes. We see there that the improvement for c_M and c_B is also better—a factor of ∼1.4—while for the rest of the cosmological parameters without priors, it is a factor of ∼1.1–1.2.

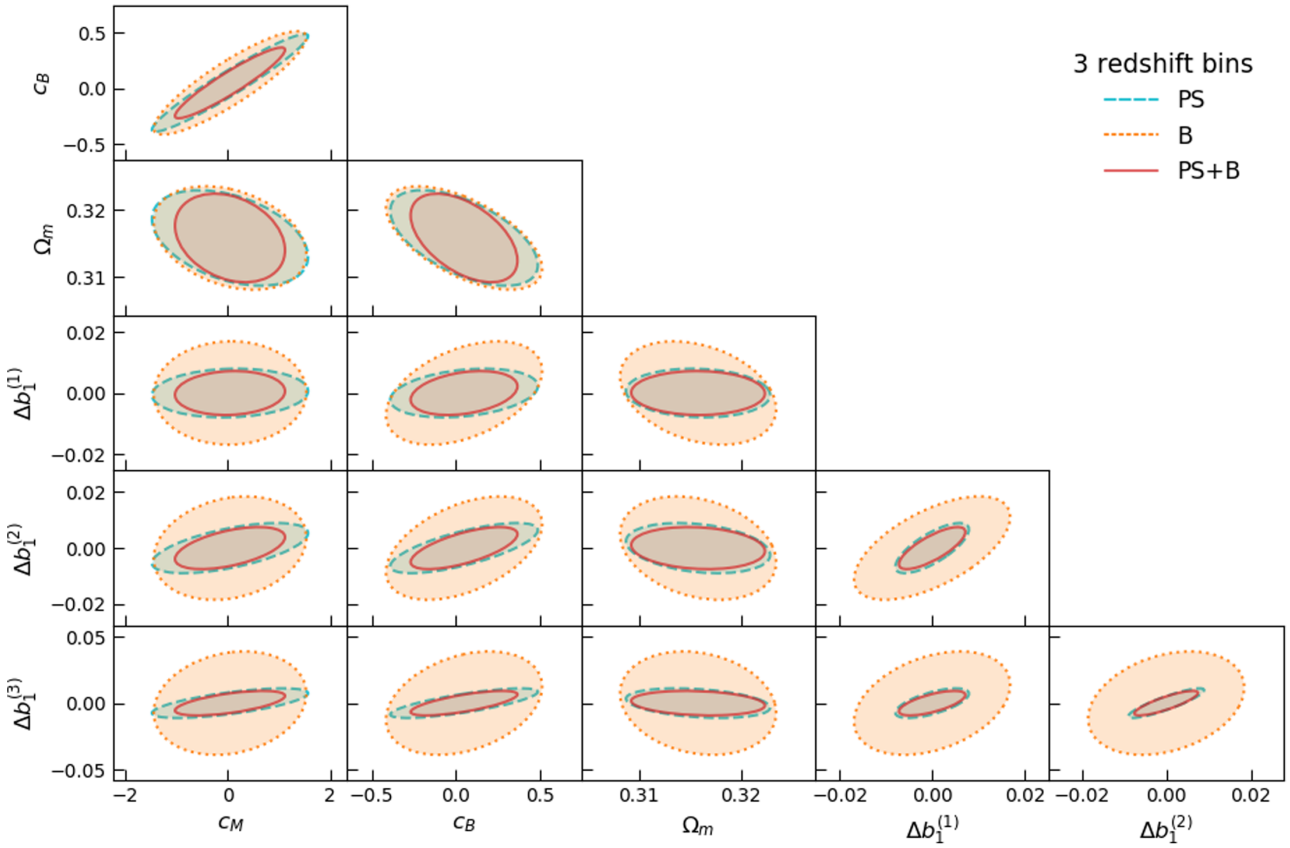

FIG. 9. Same as Fig. 8 but for $l_{\text{max}} = 3000$. Combining the power spectra and bispectra improves on the power alone by a factor of ∼1.4 for both c_M and c_B.

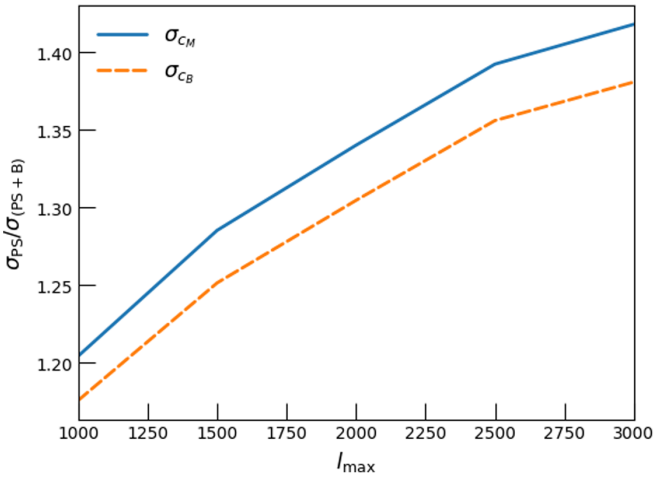

FIG. 10. Improvement on σ_{c_M} and σ_{c_B} from adding bispectra to power spectra as a function of the maximum multipole $l_{\max}$ for the $n_{\rm zbin} = 3$ case. Results are similar for other $n_{\rm zbin}$.

To see the trend more clearly, we plot in Fig. 10 the improvement from adding the bispectra on σ_{c_M} (blue solid line) and σ_{c_B} (orange dashed line) as a function of $l_{\max}$. It is clear that the higher the maximum multipole, the better the improvement gets when adding the bispectra. We also note that these forecasts of improvement at higher l are approximate, because of the nonlinear effects that become stronger at small scales that are not modeled here. However, the nonlinear effects would increase the sensitivity of both the bispectra and the power spectra, and whether the bispectra would benefit much more is to be seen. Nevertheless, the improvement due to the more rapidly growing number of modes in the bispectra would still be present.

Finally, we show in Fig. 11 how the total results vary with different numbers of redshift bins, for the fiducial choice of $l_{\max} = 1000$. It is clear that tomography serves to improve constraints: a factor of 1.5 and 1.6 for c_M and c_B, respectively, going just from 1 to 2 redshift bins; and a factor of 1.8 and 2.0 for 3 redshift bins. The 1D marginalized constraints for all the parameters without priors for $n_{\rm zbin} = 1$, 2, and 3 can be found in Table I.

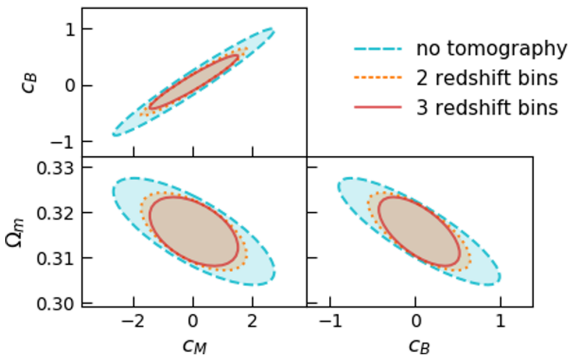

FIG. 11. Marginalized 2D constraints on parameters from a combined power spectra and bispectra analysis, varying $n_{\rm zbin} = 1$, 2, and 3 at fixed $l_{\max} = 1000$. One gains a factor of 1.8 (2.0) improvement in c_M (c_B) going from $n_{\rm zbin} = 1$ to 3, and a factor of 1.5 (1.6) from $n_{\rm zbin} = 1$ to 2.

From the way the contours shrink, it seems that the improvement will likely be marginal going much beyond $n_{\rm zbin} = 3$. This is consistent with the findings of Ref. [9], where the signal-to-noise ratio of the lensing convergence bispectrum plateaus for $n_{\rm zbin} \geq 5$ with or without the non-Gaussian and supersample covariance. Given that the lensing bispectrum is the main beneficiary of tomography, we expect similar conclusions to hold for our combined bispectrum results, and so we do not explore much higher numbers of redshift bins. Note also that the bispectrum computation becomes expensive for high $n_{\rm zbin}$ as the number of unique tomographic combinations increases quickly with more bins.

VII. SUMMARY AND DISCUSSION

In this paper we forecasted the ability of the Roman Space Telescope overlapped with the LSST survey over its notional 2000 deg^2 survey to constrain a subclass of Horndeski theories, by using galaxy and lensing convergence bispectra in addition to power spectra. In particular, we explored the cross-bispectra as a way to improve constraints over any auto-bispectra alone. We summarize the main results below. They are quoted for our fiducial choice of $l_{\max} = 1000$ and three redshift bins unless otherwise stated. Note that we have also left off the treatment of systematic effects such as intrinsic alignments, baryonic effects, and photometric redshift, which should be evaluated in future work:

(i) Combining all possible auto- and cross-bispectra between the two types of probes κ and g gave factors of 1.6 and 1.5 better constraints on c_M and c_B, respectively, compared to using $\kappa\kappa\kappa$ type of bispectra alone, and a factor of 8 and 52 compared to using ggg type alone.

(ii) Including all possible cross-bispectra between κ and g in different tomographic bins contributed to factors of 1.3 and 1.1 improvement on the c_M and c_B constraints, respectively, compared to using all the auto-bispectra defined as

$$\{\kappa_{(i)}\kappa_{(i)}\kappa_{(i)}, g_{(i)}g_{(i)}g_{(i)} | i = 1, 2 \cdots n_{\rm zbin}\}.$$

(iii) Adding the combined bispectrum result to the power spectrum led to a factor of ∼1.2 improvement for both MG parameters, yielding $\sigma_{c_M} = 1.0$ and $\sigma_{c_B} = 0.3$.

(iv) Varying the $l_{\max}$ used, we find that the improvement due to bispectra increases to a factor of ∼1.4 for both MG parameters for $l_{\max} = 3000$. This is primarily due to the greater number of modes in bispectrum with $l_{\max}$. While we expect our linear modeling to be

less accurate (in fact, more conservative) at increasingly $l_{\rm max}$ and that the absolute values of the constraints would change with nonlinear modeling added, we expect the relative improvement factor to remain similar.

(v) Varying $n_{\rm zbin} = 1$, 2, and 3, we find that using two tomographic bins already gives factors of 1.5 and 1.6 improvement in the c_M and c_B constraints compared to no tomography; whereas using three bins leads to a factor of 1.8 and 2.0 better constraints. For the Roman + LSST survey considered here, the improvement beyond three redshift bins is likely to be marginal, while it also becomes computationally expensive to go to higher $n_{\rm zbin}$ as the number of bispectra combinations scales rapidly with $n_{\rm zbin}$.

We caution the readers that these results were obtained using modeling that is solely valid in the linear regime, while the observables are integrated along the line of sight and in principle capture scales down to the nonlinear regime. While we varied one $l_{\rm max}$ as a very crude way to control the impact of nonlinear scales, there exists more refined methods such as using a different $l_{\rm max}$ per redshift bin corresponding to a desired $k_{\rm max}$, e.g., Ref. [18].

Another less explored method but highly relevant for data analysis is to cut out actual physical scales by forming the appropriate linear combinations of the observables, by extending the k-cut method originally proposed in Ref. [31] for C_l's. As the nonlinear modeling of MG theories are still underway (templates have been suggested and partially tested on equilateral triangles for a few modified gravity models in Ref. [29]), a k-cut method for bispectra could help to control the exact $k_{\rm max}$ allowed by the modeling available at the time of data analysis.

Regardless of the caveats named above, we expect that the general observation that cross-bispectra could be powerful at breaking parameter degeneracy to remain applicable to many cases and extendable to other experiments as well. Therefore, this work opens the way for combining multiple probes in higher-order statistics, and providing more avenues for maximizing the information content of next-generation large-scale-structure surveys.

ACKNOWLEDGMENTS

We thank Masahiro Takada, Bhuvnesh Jain, Wayne Hu, Tim Eifler, Atsushi Taruya, Ben Bose, Hiroyuki Tashiro, Hayato Motohashi, Zachary Slepian, Kris Pardo, and Agnes Ferté for useful discussions. We thank the Nancy Grace Roman Space Telescope *Cosmology with the High Latitude Survey* Science Investigation Team for providing feedback for this work and the redshift distributions used in the forecast. C. H. especially thanks Miguel Zumalacárregui for providing private versions of `hi_class` and guidance on using the software; C. H. also thanks Tessa Baker for sharing her personal notes on Horndeski theories. Part of this work was done at Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. California Institute of Technology. Government sponsorship acknowledged.

APPENDIX A: SECOND ORDER EXPANSION OF WIGNER-3j SYMBOLS WITH STIRLING APPROXIMATION

Calculating the bispectrum involves evaluating the Wigner-3j symbol, which has a closed algebraic form (e.g., Hu 2000):

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} = (-1)^L \frac{L!}{(L-l_1)!(L-l_2)!(L-l_3)!} \times \left[\frac{(2L-2l_1)!(2L-2l_2)!(2L-2l_3)!}{(2L+1)!}\right]^{1/2} \tag{A1}$$

for even $l_1 + l_2 + l_3$ and zero for odd $l_1 + l_2 + l_3$, where we have also defined $L = (l_1 + l_2 + l_3)/2$.

Because evaluating the exact expression involves calculating factorials $l!$ which diverge for large l, we employ an approximation based on the Stirling approximation: $n! = \Gamma(n+1)$ and

$$\Gamma(x) \sim (2\pi)^{1/2} e^{-x} x^{x-1/2}, \quad \text{for large } x. \tag{A2}$$

While the commonly used first order expansion is good enough with subpercent errors for angular scales $l \geq 50$

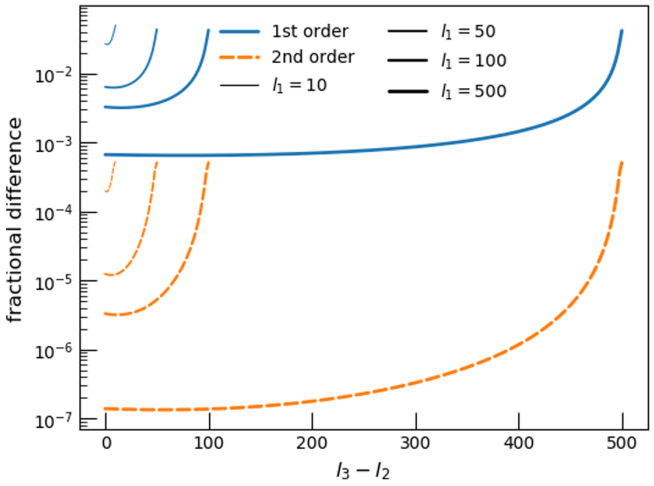

FIG. 12. Fractional difference between the Stirling approximation and the exact computation of the Wigner-3j symbols as a function of $l_3 - l_2$ for fixed $l_1 = l_2$. The commonly used first order approximation (solid lines) is known to reach subpercent level accuracy on scales of interest $l \geq 50$ except for degenerate triangles (e.g., Ref. [8]). The second order approximation (dashed lines), on the other hand, requires negligible additional computation but reduces the error on degenerate triangles by more than 1 order of magnitude, reaching subpercent level accuracy for all configurations.

of our interest for most triangles, it is known to be less accurate for the degenerate triangles, with errors reaching sometimes above percent level (see Fig. 12 for an illustration). We therefore expand the expression to second order and found that it reduces the error by more than an order of magnitude for degenerate triangles, rendering it subpercent. At the same time, the errors on other configurations are in general $< 10^{-5}$. These improvements are obtained with negligible additional computational cost, and we recommend using them for any bispectrum calculations.

The approximate expression is given by

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \approx (-1)^L \sqrt{\frac{e^3}{2\pi}} (L+1)^{-1/4} \left[\prod_{i=1}^{3} (L-l_i+1)^{-1/4} \left(\frac{L-l_i+1/2}{L-l_i+1} \right)^{L-l_i+1/4} \right] F(l_1,l_2,l_3), \tag{A3}$$

where $F(l_1,l_2,l_3)=1$ for the commonly used first order expansion, and

$$F(l_1,l_2,l_3) = \left(1+\frac{1}{12(L+1)}\right)\left(1+\frac{1}{24(L+1)}\right)^{-1/2} \prod_{i=1}^{3} \left(1+\frac{1}{12(L-l_i+1)}\right)^{-1} \left(1+\frac{1}{24(L-l_i+1/2)}\right)^{1/2} \tag{A4}$$

for the second order expansion used in this paper.

APPENDIX B: THE EXPRESSIONS FOR $\lambda(a)$ IN F_2 FOR $\alpha_i(a)=c_i\Omega_{\rm DE}(a)$ HORNDESKI THEORIES

We model the modified gravity effects in the bispectrum up to second order in perturbation theory, where the second order kernel F_2 is modified with a parameter $\lambda(a)$. In Sec. IV B we introduced the ansatz $\lambda(a) \approx \tilde{\Omega}_m^{\xi}(a)$. We now summarize briefly the first order expansion of ξ in terms of c_i's that we used to compute $\lambda(a)$ in this work, and we refer the readers to Ref. [20] for more details.

For $\alpha_M(a) = c_M(1-\tilde{\Omega}_{\rm m}(a))$, we can expand ξ in orders of $\Omega_{\rm DE}$ where the leading order approximation is given by

$$\xi = \frac{-3+6\tilde{\gamma}+2\kappa_\Phi^{(1)}+7\tau_\Phi^{(1)}}{(7-6w^{(0)}+2c_{\rm M})(1-3w^{(0)}+c_{\rm M})}, \tag{B1}$$

where

$$\tilde{\gamma} \approx \frac{3-3w^{(0)}-2\kappa_\Phi^{(1)}}{5-6w^{(0)}+2c_{\rm M}} \tag{B2}$$

and $w^{(0)}$, $\kappa_\Phi^{(1)}$, and $\tau_\Phi^{(1)}$ are lowest order coefficients in the expansion of the dark energy equation of state $w_{\rm DE}=p_{\rm DE}/\rho_{\rm DE}$, κ_Φ, and τ_Φ, respectively,

$$w_{\rm DE} = \sum_{n=0} \frac{1}{n!} w^{(n)} (1-\tilde{\Omega}_{\rm m})^n, \tag{B3}$$

$$\kappa_\Phi - \frac{3}{2}\tilde{\Omega}_{\rm m} = \sum_{n=1}^{\infty} \frac{1}{n!} \kappa_\Phi^{(n)} (1-\tilde{\Omega}_{\rm m})^n, \tag{B4}$$

$$\tau_\Phi = \sum_{n=1}^{\infty} \frac{1}{n!} \tau_\Phi^{(n)} (1-\tilde{\Omega}_{\rm m})^n. \tag{B5}$$

They can be written in terms of c_i's in theories where $\alpha_i(a) = c_i(1-\tilde{\Omega}_{\rm m}(a))$ as

$$\kappa_\Phi^{(1)} = \frac{3}{2}\left[c_{\rm T} + \frac{(c_{\rm B}+2c_{\rm M}-2c_{\rm T})^2}{6(1+w^{(0)})-6c_{\rm B}w^{(0)}+2c_{\rm B}c_{\rm M}-c_{\rm B}+4c_{\rm M}-4c_{\rm T}} \right], \tag{B6}$$

$$\begin{aligned} \tau_\Phi^{(1)} = {} & \frac{1}{3}\left(\kappa_Q^{(0)}\right)^2 \left[\left(c_{\rm B}+c_{\rm M}-\frac{3}{2}c_{\rm T} \right)\kappa_Q^{(0)} - \frac{9}{4}c_T \right] + \frac{1}{3}\left(\kappa_Q^{(0)}\right)^2 \left[(1+3w^{(0)}-c_{\rm M})\kappa_Q^{(0)} - \frac{9}{2} \right] c_{\rm V1} \\ & - \frac{1}{2}\kappa_Q^{(0)} \left(\kappa_Q^{(0)} + \frac{3}{2} \right) \left[\left(\frac{1}{2} - 3w^{(0)} + c_{\rm M} \right)\kappa_Q^{(0)} + 3 \right] c_{\rm V2}, \end{aligned} \tag{B7}$$

where

$$\kappa_Q^{(0)} = -\frac{3(c_{\rm B}+2c_{\rm M}-2c_{\rm T})}{6(1+w^{(0)})-6c_{\rm B}w^{(0)}+2c_{\rm B}c_{\rm M}-c_{\rm B}+4c_{\rm M}-4c_{\rm T}}. \tag{B8}$$

With these expressions ξ can be expressed completely in terms of the constant parameters $\{w^{(0)}, c_{\rm B}, c_{\rm M}, c_{\rm T}, c_{\rm V1}, c_{\rm V2}\}$. We assume ΛCDM cosmology as the background expansion so that $w^{(0)} = -1$ throughout. Note that $c_{\rm V1}$ and $c_{\rm V2}$ are constants of proportionality for the functions α_{V1} and α_{V2}. While $\alpha_i, i = M, B, K, T$ describe the first order degrees of freedoms of the Horndeski Lagrangian, α_{V1} and α_{V2} are part of the second order expansion, whose relations to G_a are given by

$$M^2\alpha_{\rm V1} = -2X(G_{4X} + 2XG_{4XX} - G_{4\phi} + 2H\dot{\phi}G_{5X} - XG_{5\phi X} + \dot{\phi}HXG_{5XX}), \tag{B9}$$

$$M^2\alpha_{\rm V2} = 2\dot{\phi}HXG_{5X}. \tag{B10}$$

Although they are in principle arbitrary functions, we restrict ourselves to setting $c_{V1} = c_{V2} = 0$ when evaluating F_2.

Note that we have assumed that the constant ξ ansatz is a good approximation to the models used here. The validity of the approximation is checked explicitly against solving for $\lambda(a)$ numerically for various choices of c_i's in Ref. [20]. For most cases the deviation is less than 10% while in some cases the constancy is eventually violated at low redshifts (e.g., $c_B = 0.2, c_T = 0.5$). We expect that for the much smaller values of c_i's considered here, the deviations would be less significant.

[1] R. Laureijs *et al.*, arXiv:1110.3193.
[2] D. Spergel *et al.*, arXiv:1503.03757.
[3] v. Z. Ivezić *et al.* (LSST Collaboration), Astrophys. J. **873**, 111 (2019).
[4] A. Aghamousa *et al.* (DESI Collaboration), arXiv:1611.00036.
[5] T. Eifler *et al.*, arXiv:2004.05271.
[6] T. Eifler *et al.*, arXiv:2004.04702.
[7] S. Hemmati, P. Capak, D. Masters, I. Davidzon, O. Dorè, J. Kruk, B. Mobasher, J. Rhodes, D. Scolnic, and D. Stern, Astrophys. J. **877**, 117 (2019).
[8] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **348**, 897 (2004).
[9] M. Rizzato, K. Benabed, F. Bernardeau, and F. Lacasa, Mon. Not. R. Astron. Soc. **490**, 4688 (2019).
[10] V. Yankelevich and C. Porciani, Mon. Not. R. Astron. Soc. **483**, 2078 (2019).
[11] M. Tellarini, A. J. Ross, G. Tasinato, and D. Wands, J. Cosmol. Astropart. Phys. 06 (2016) 014.
[12] G. W. Horndeski, Int. J. Theor. Phys. **10**, 363 (1974).
[13] E. Bellini and I. Sawicki, J. Cosmol. Astropart. Phys. 07 (2014) 050.
[14] D. Bettoni and M. Zumalacárregui, Phys. Rev. D **91**, 104009 (2015).
[15] M. Zumalacárregui, E. Bellini, I. Sawicki, J. Lesgourgues, and P. G. Ferreira, J. Cosmol. Astropart. Phys. 08 (2017) 019.
[16] J. M. Ezquiaga and M. Zumalacárregui, Phys. Rev. Lett. **119**, 251304 (2017).
[17] R. Kase and S. Tsujikawa, Int. J. Mod. Phys. D **28**, 1942005 (2019).
[18] D. Alonso, E. Bellini, P. G. Ferreira, and M. Zumalacárregui, Phys. Rev. D **95**, 063502 (2017).
[19] J. Gleyzes, D. Langlois, M. Mancarella, and F. Vernizzi, J. Cosmol. Astropart. Phys. 02 (2016) 056.
[20] D. Yamauchi, S. Yokoyama, and H. Tashiro, Phys. Rev. D **96**, 123516 (2017).
[21] E. Bellini, I. Sawicki, and M. Zumalacárregui, J. Cosmol. Astropart. Phys. 02 (2020) 008.
[22] J. Lesgourgues, arXiv:1104.2932.
[23] C. M. Hirata, N. Gehrels, J.-P. Kneib, J. Kruk, J. Rhodes, Y. Wang, and J. Zoubian, arXiv:1204.5151.
[24] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **395**, 2065 (2009).
[25] M. Ishak *et al.*, arXiv:1905.09687.
[26] Y. Takushima, A. Terukina, and K. Yamamoto, Phys. Rev. D **92**, 104033 (2015).
[27] Y. Takushima, A. Terukina, and K. Yamamoto, Phys. Rev. D **89**, 104007 (2014).
[28] R. Scoccimarro and H. M. P. Couchman, Mon. Not. R. Astron. Soc. **325**, 1312 (2001).
[29] B. Bose, J. Byun, F. Lacasa, A. Moradinezhad Dizgah, and L. Lombriser, J. Cosmol. Astropart. Phys. 02 (2020) 025.
[30] T. Lazeyras, C. Wagner, T. Baldauf, and F. Schmidt, J. Cosmol. Astropart. Phys. 02 (2016) 018.
[31] P. L. Taylor, F. Bernardeau, and T. D. Kitching, Phys. Rev. D **98**, 083514 (2018).